



# Le haka des All Blacks : une expression vocale complexe !

Michel Chafcouloff

## ► To cite this version:

Michel Chafcouloff. Le haka des All Blacks : une expression vocale complexe!. Travaux Interdisciplinaires sur la Parole et le Langage, 2005, 24, pp.13-84. hal-00241550

**HAL Id: hal-00241550**

**<https://hal.science/hal-00241550>**

Submitted on 6 Feb 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# LE HAKA DES ALL BLACKS : UNE EXPRESSION VOCALE COMPLEXE !

Michel Chafcouloff

## Résumé

*Suite à une étude préliminaire concernant la relation entre voix et gestes dans l'exécution du Haka par les joueurs de rugby néo-zélandais, la composante vocale a été l'objet d'une étude afin de spécifier le type de voix produite par les All Blacks, c'est-à-dire une voix criée ou une voix chantée. Le signal de parole (bruitée) a été filtré, numérisé et l'analyse acoustique a été fondée sur l'utilisation des logiciels Phonedit et Praat. L'examen des spectrogrammes de parole montre que la durée des voyelles est largement supérieure à celle des consonnes (80-20%), qu'une élévation de la fréquence fondamentale jusqu'à 400 Hz entraîne des variations importantes dans les fréquences de formants des voyelles et que l'intensité mesurée ne rend pas compte de l'intensité réelle. En ce qui concerne le rythme, on distingue deux types de relations entre le rythme gestuel et le rythme musical, et le rythme gestuel et le rythme verbal. En bref, il ressort que la voix produite par le chœur des All Blacks présente les attributs acoustiques de la voix criée. Toutefois, compte tenu de son organisation temporelle et de sa structure rythmique, le Haka peut être également considéré comme une composition musicale, et de ce fait assimilé à un chant.*

Mots-clés : acoustique, chant maori, cri, para-linguistique, relation voix/gestes.

## Abstract

*Following a preliminary study concerning the relationship between voice and gestures in the performance of the Haka by the rugby players of New Zealand, the vocal component has been subjected to a study to specify the type of voice produced by the All Blacks, i.e. a shouting voice or a singing voice. The (noisy) sound signal was filtered, digitized and the acoustic analysis was carried out using the Phonedit and Praat softwares. Examination of spectrograms showed that the duration of vowels is considerably longer than that of consonants (80% vs 20%), that a rise of F0 up to 400 Hz brings about drastic changes in vowel formant frequencies and that there is a discrepancy between 'measured' and 'real' intensity. As for rhythm, two types of relationships should be distinguished: the relationship between gestural rhythm and musical rhythm, and the relationship between gestural rhythm and verbal rhythm. In brief, it may be concluded that the voice produced by the All Blacks chorus has the acoustical attributes of a 'shouting voice'. Nevertheless, if attention is paid to its temporal organisation and rhythmical structure, the Haka should also be considered as a musical composition, and hence as a chant.*

Keywords : acoustics, para-linguistics, voice/gesture relationship, shouting voice.

---

CHAFCOULOFF, Michel (2005), *Le Haka des All Blacks : une expression vocale complexe !*, Travaux Interdisciplinaires du Laboratoire Parole et Langage, vol. 24, p. 13-83.

*I understand a fury in your words,  
but not the words*  
Shakespeare, Othello, Act. 4

## Avant-propos

Toute recherche scientifique est rarement le fruit d'une réflexion exclusivement personnelle. Le plus souvent, elle émane de discussions ou d'échanges de vues qu'a pu avoir le chercheur avec des collègues ou amis qui partagent le même intérêt ou la même passion, discussions qui serviront éventuellement de prélude à une étude collective ou individuelle. Au terme de celle-ci, il nous a semblé opportun d'exposer brièvement comment et pourquoi nous avons été amené à nous intéresser au Haka, sujet dont le moins que l'on puisse dire, est qu'il se situait aux antipodes (la comparaison est toute trouvée) des thèmes de recherche qui avaient été les nôtres jusqu'alors.

Tout a commencé par des discussions enflammées dans les couloirs de notre laboratoire, au cours desquelles il nous arrivait de commenter avec Jacques Boyer, Toulonnais de pure souche comme l'auteur de la présente étude, les résultats des matches du Rugby Club Toulonnais. Cette passion commune pour le R.C.T. où le chauvinisme le plus partisan était la règle, nous a entraînés à discuter du rugby néo-zélandais et d'une équipe maorie que nous avons eu l'occasion de voir au stade Mayol à Toulon à l'occasion d'une tournée. Le lien entre la passion sportive et l'intérêt scientifique était évident et n'a pas tardé à se concrétiser. Dans le cadre des études qui étaient conduites concernant les relations entre Oralité et Gestualité au sein de notre laboratoire, pourquoi ne pas étudier la production du message oral en relation avec le comportement gestuel pendant l'exécution du Haka, le fameux cri de guerre poussé par les All Blacks. L'idée était lancée, les locuteurs et le corpus étaient trouvés, il ne restait plus qu'à se mettre au travail, et les premiers résultats firent l'objet d'une communication (Chafcouloff *et al.*, 2001) à l'occasion du colloque ORAGE 2001, dans laquelle il était démontré qu'il existait une relation entre gestes et voix, c'est-à-dire que des claquements résultant du contact des mains avec différentes parties du corps étaient synchronisés avec la production d'unités phoniques spécifiques, en l'occurrence les occlusives sourdes pour donner au Haka le rythme de base qui lui est propre. Deux années plus tard, nous donnions à ce travail une audience plus large à l'occasion d'une deuxième communication présentée au XV<sup>e</sup> Congrès International des Sciences Phonétiques à Barcelone (2003). Toutefois,

l'étude de la relation entre les composantes vocale et gestuelle ne constituait qu'une étape et d'autres points d'intérêt restaient à élucider. C'est ainsi que nous nous sommes attaché à définir le degré de correspondance entre graphisme et phonétisme, et à décrire la substance phonique sur le plan acoustique. Ensuite, il nous appartenait d'examiner la structure rythmique du Haka et de préciser le type de voix produite par le chœur des All Blacks. L'extension de l'étude s'est révélée source d'enrichissement, puisque nous avons été amené à nous intéresser à une langue peu connue, en l'occurrence le maori, mais à traiter également des questions aussi diverses que l'acoustique de la voix criée et de la voix chantée, la psychologie du rythme, la symbolisation subjective des unités phoniques, la rhétorique ou la poétique, et d'autres domaines encore où nos connaissances étaient insuffisantes, voire inexistantes.

De nombreuses personnes nous ont apporté le témoignage de leur intérêt pour ce travail qui n'aurait pas été mené à bien sans leur concours, et il nous appartient aujourd'hui de les remercier pour leur aide.

En premier lieu, je voudrais remercier les personnes que nous avons contactées en Nouvelle-Zélande et qui nous ont répondu spontanément, surprises qu'elles étaient qu'un chercheur français puisse faire du *Haka* un sujet d'étude. En particulier, Margaret Mac Lagan de l'université de Canterbury et P. Warren de l'université de Wellington qui nous ont fait parvenir une documentation fort précieuse sur la langue maorie. Également Paul Waite pour son article sur l'histoire du *Haka* via Internet ; grâce à lui, nous avons découvert l'ouvrage de Timoti Karetu, *Haka: the Dance of a noble People*, lequel nous a d'ailleurs aimablement répondu, et ce livre est rapidement devenu notre ouvrage de référence. Enfin, Andrew Martin, manager général d'une tournée des All Blacks en 2002, qui le premier, nous a fait part de son intérêt pour cette étude. Que tous mes correspondants et nouveaux amis néo-zélandais soient assurés en cette occasion de ma profonde gratitude pour leur aide spontanée.

Sur le plan local, je voudrais remercier tout d'abord J. Boyer et I. Guaitella avec lesquels j'ai débuté cette étude, le professeur Mario Rossi pour les entretiens que nous avons eus, le personnel technique C. Deniaud, A. Ghio et L. Seimandi dont l'assistance dans le domaine de l'informatique a été grandement appréciée. Mes collègues du laboratoire Parole et Langage qui ont bravement enduré le 'supplice' des tests de perception au cours desquels leurs oreilles étaient soumises à rude épreuve à l'écoute des cris gutturaux poussés par les All Blacks. En particulier, les relecteurs, qui chacun dans un domaine particulier, nous ont fait bénéficier de leurs remarques autant pertinentes que constructives : A. Di Cristo dans le domaine de la prosodie, notre collègue polonais R. Gubrynowicz en acoustique, D. Autesserre dans le domaine de l'anthropologie, N. Scotto di Carlo

dont les compétences unanimement reconnues dans le domaine de l'odologie et de la voix chantée nous ont été infiniment précieuses, ainsi que D.J. Hirst pour sa disponibilité jamais démentie.

Enfin, mes remerciements vont tout naturellement aux All Blacks pour m'avoir donné l'occasion de travailler sur un sujet 'original' de l'aveu même des personnes auxquelles nous faisons part de nos orientations de recherche, sujet qui m'a passionné car il me permettait de concilier intérêt scientifique et passion sportive. Quel chercheur n'a pas souhaité travailler en s'amusant ou *vice versa*. Tel a été mon cas, et je mesure pleinement la chance que j'ai eue ces dernières années. Mon intérêt pour les All Blacks n'est pas nouveau et remonte en fait à 1967 quand j'ai découvert à l'occasion de leur tournée en Europe, la génération des B. Lochore, C. Meads, W. Nathan, C. Laidlaw et autres, marée noire qui déferlait sur les terrains de rugby du vieux continent et qui emportait tout sur son passage. L'intérêt est devenu rapidement passion, laquelle ne m'a jamais quitté, essentiellement pour une certaine façon de pratiquer le rugby '*ce jeu de voyous pratiqué par des gentlemen*' ainsi que pour une philosophie propre aux hommes en noir pour qui '*le rugby n'est jamais que la meilleure façon de bien vivre une jeunesse*' pour reprendre en substance les termes de Zinzan Brooke, ancien capitaine des Blacks. Depuis lors et chaque fois que j'en ai eu l'occasion, j'ai suivi avec ferveur les matches des générations des Blacks qui ont suivi, joyeux avec eux les soirs de victoires et morose les soirs de leurs (rares) défaites.

Merci aux Blacks des joies qu'ils nous donnent, et d'être ce qu'ils ont toujours été, c'est-à-dire LA référence en matière de rugby ; un groupe, une équipe qui, par son éthique sportive, a su s'attirer le respect unanime dans le monde de 'l'Ovalie', et qui fait que dans l'esprit et dans le cœur des nombreux amateurs de ce sport, il y a un 'black fan' qui sommeille. À mon modeste niveau, peut-être aurai-je contribué sur le plan scientifique ainsi que sur celui de la connaissance générale à mieux faire connaître et surtout respecter le *Haka*. Si tel est le cas, alors mon but aura été atteint.

*Aux hommes en noir, à ceux qui, toutes générations confondues,  
ont su rester 'bumbles' dans le succès comme 'dignes' dans le revers.  
La marque des grands.*

## Introduction

Dans certaines langues du monde, il est parfois difficile d'établir une distinction stricte et non ambiguë entre les deux principales formes d'expression orale que sont le chant et la parole. En effet, si l'on considère les différences qui existent entre les langues tant sur le plan segmental que suprasegmental, on peut formuler l'hypothèse qu'une langue qui serait caractérisée par la richesse de son système vocalique, et dont les voyelles se trouveraient, de surcroît, affectées de variations mélodiques de forte amplitude, sera selon toute probabilité perçue comme plus mélodieuse qu'une autre langue comparativement pauvre dans les deux domaines précités. Dans le premier cas, il est vraisemblable que la voix du locuteur sera plutôt assimilée à une voix chantée, alors que dans le second cas, elle sera perçue comme une voix parlée. De plus, si on tient compte des caractéristiques idiosyncrasiques des locuteurs, et que l'on considère que, indépendamment des langues ou des dialectes, la voix de certains d'entre eux se distingue par un 'accent naturellement chantant', il semble que la parole puisse être assimilée à une mélodie phrastique et soit perçue comme un chant, si les tenues vocaliques atteignent une durée critique (Scotto di Carlo, 1994) ou si la parole dépasse un certain degré de musicalité, cette dernière ayant été définie par Fónagy (1983) comme *la régularité de la distribution des fréquences fondamentales à l'intérieur d'une syllabe* (Fónagy, *op. cit.*, p. 310).

L'existence de plusieurs types de parole différant en termes de degrés de musicalité est attestée dans la langue maorie de Nouvelle-Zélande, langue dans laquelle on trouve entre les deux catégories extrêmes que sont la parole *koorero* et le plain-chant *waiata*, deux catégories intermédiaires, le *karakia* assimilé à une mélopée rituelle, et le *Haka* qui correspond pour sa part à une parole "stylisée et mélodieuse" (Fónagy, *op. cit.*, p. 311). En ce qui concerne ce dernier type de parole, il convient de préciser que, si le mot 'Haka' est connu des linguistes versés dans l'étude des langues polynésiennes qui l'associent effectivement à un type de parole particulier, le terme générique de 'Haka' évoque pour le grand public et en particulier pour les amateurs de rugby, une danse qui était exécutée aux temps des guerres tribales par les guerriers maoris dans le but d'impressionner leurs ennemis et montrer ainsi leur détermination avant la bataille. Cette danse, où voix et gestes sont coordonnés, fût adoptée vers la fin du XIX<sup>e</sup> siècle par les joueurs de rugby de Nouvelle-Zélande, plus connus sous le nom de All Blacks (ABs) du fait qu'ils étaient vêtus de noir, et fût exécutée pour la première fois en Europe, à l'occasion d'une tournée en Grande-Bretagne en 1905. Depuis lors, le *Haka* est devenu un rituel d'avant-match incontournable, et est connu des amateurs de ce sport grâce aux retransmissions télévisées des tournées des All Blacks en Europe, mais aussi à l'occasion de la Coupe du Monde de Rugby et du tournoi des 'Tri Nations', où sont opposées les trois nations majeures de l'hémisphère sud (Afrique du Sud, Australie et Nouvelle-Zélande).

Dans des études préliminaires consacrées à la relation entre voix et gestes (Chafcouloff *et al.*, 2001) ; Chafcouloff, 2003), nous avons démontré qu'il existait une synchronisation temporelle entre la composante vocale et la composante gestuelle dans l'exécution du *Haka*, en ce sens que certaines syllabes, comprenant des consonnes occlusives en position initiale, sont produites en synchronisation avec des claquements résultant du contact des mains des joueurs avec différentes parties du corps (cuisses, poitrine, bras opposé), etc.

Dans la présente étude qui s'inscrit comme la suite logique des précédentes, notre attention a été concentrée plus particulièrement sur le message verbal transmis au cours du *Haka*. Dans ce but, une analyse acoustique fine du signal de parole a été effectuée afin de répondre à plusieurs questions restées sans réponse au terme des études préliminaires :

- dans quelle mesure la version orale du *Haka* exécuté et interprété par les All Blacks correspond-elle à la version écrite du texte original ? En effet, si l'on se fonde sur l'opinion exprimée par Karetu (1993, p. 68), selon lequel les mots sont souvent déformés ou mal prononcés, ([...] *the words are often bowdlerised and mispronounced*), il semble que d'importantes modifications soient susceptibles d'affecter la structure originelle du message verbal. En d'autres termes, quel est le degré de correspondance entre graphisme et phonétisme ?
- sur le plan segmental, quelles sont les unités phoniques qui se trouvent être privilégiées et selon quelle distribution ? La tension psychologique qui prélude à l'affrontement physique entre deux équipes de rugby entraîne-t-elle des changements dans la structure acoustique des unités phoniques en relation avec le fonctionnement particulier de la source sonore ?
- sur le plan suprasegmental, la voix émanant du chœur des All Blacks est-elle caractérisée par des variations mélodiques conséquentes, et dans quel registre tonal se situent-elles ? D'autre part, existe-t-il une structure rythmique propre au *Haka*, et quels sont le ou les paramètres acoustiques en rapport avec cette rythmicité ? Quelle est l'organisation temporelle de la syllabe en termes de rapports de durée entre les voyelles et les consonnes ? Existe-t-il une relation entre rythme gestuel et rythme verbal ? Comment les prééminences accentuelles sont-elles perçues ? Enfin, quel est le rôle des pauses ?

Brièvement exposées, telles sont les principales questions auxquelles il nous appartenait de répondre, afin de déterminer si le message verbal transmis par les All Blacks correspondait (ou non) à la définition du *Haka* telle qu'elle en a été donnée par Fónagy (*op. cit.*). Puisqu'il existe plusieurs types de *Haka* qui diffèrent tant du point de vue des thèmes qui y sont abordés que de l'expression vocale et gestuelle qui les caractérisent (Karetu, 1993), quel est le type de parole

approprié pour décrire le *Haka* exécuté par les All Blacks ? Plutôt qu'une parole 'stylisée et mélodieuse', pour reprendre les termes employés par Fónagy, le *Haka* ne serait-il pas plutôt une parole 'scandée et martelée' ? Dans ce cas, l'emploi du terme 'mélodieux' est-il justifié sur le plan acoustique et perceptif ? Enfin, pour répondre à la question que nous posons dans le titre et qui a trait à la complexité du message vocal, dans quelle mesure la voix produite par le chœur des All Blacks doit-elle être assimilée à une voix 'parlée', 'criée' ou 'chantée' ?

### 1.1. Les différents types de *Haka*

Comme nous l'avons mentionné plus haut, le terme de *Haka* évoquait à l'origine une danse qui était exécutée par les guerriers maoris avant la bataille. Dans le livre de T. Karetu (*op. cit.*), *Haka! The Dance of a Noble People*, il est précisé qu'il existe en réalité non pas un, mais plusieurs types de *Haka*. Dans son ouvrage, l'auteur souligne les spécificités de chacun d'entre eux, et décrit l'art d'exécuter le *Haka* tant sur le plan gestuel que vocal. Il insiste sur l'importance que celui-ci revêt pour les maoris, n'en serait-ce pour preuve l'organisation de concours de *Haka*, où sont jugés non seulement l'harmonie vocale, mais également le comportement gestuel et l'expression faciale, c'est-à-dire la mimo-gestualité des exécutants.

Le premier type de *Haka* est le '*Haka peruperu*' du mot '*peru*' (colère en maori), qui est considéré comme l'archétype de la danse guerrière. Il est caractérisé sur le plan gestuel par le maniement d'armes (lances, haches, mousquets) ainsi que par un saut vertical qui met un terme à la danse. Ce saut est accompagné par une élévation des bras et par une flexion des jambes, où les mollets entrent en contact avec la partie postérieure des cuisses. Les guerres tribales appartenant à un passé révolu, ce type de *Haka* n'est plus guère exécuté aujourd'hui qu'à l'occasion de spectacles folkloriques ou de cérémonies locales.

Plus pratiqué de nos jours est le '*Haka taparahi*' exécuté sans armes, et dont les paroles sont assurément moins belliqueuses que dans le *Haka* précédent. On y traite souvent du thème de l'intégration raciale, et en particulier des problèmes auxquels se trouvent confrontés les maoris en tant que minorité ethnique. Au-delà des revendications (légitimes) qui visent à une meilleure intégration et à la promotion sociale des maoris, l'accent est mis sur la nécessité de préserver les traditions et les valeurs ancestrales de la race. Toutefois, il convient de reconnaître que, à l'exception des néo-zélandais de souche, ainsi que des (rares) touristes étrangers *pākehā* en l'honneur desquels ce *Haka* est exécuté, il demeure généralement méconnu du grand public. Par contre, s'il existe un *Haka* dont on peut dire qu'il jouit d'une notoriété quasi universelle essentiellement grâce à la télévision, c'est le célèbre rituel '*Ka mate*' exécuté par les All Blacks avant chaque match international de rugby.



## 1.2. L'origine du *Ka mate*

Le *Ka mate* remonte aux années 1820 et se rapporte à un épisode des guerres tribales qui opposaient certaines tribus de Nouvelle-Zélande. Au cours de l'une de ces guerres, *Te Rauparaha*, le chef de la tribu des *Ngati Toa* dû s'enfuir devant ses ennemis, et chercha refuge auprès du chef d'une tribu voisine nommé *Te Wharerangi*. Non sans une certaine réticence de sa part, ce dernier l'autorisa à se cacher dans une fosse et ordonna à sa femme *Te Rangikoa* de s'asseoir sur ladite fosse, afin de le soustraire à la vue de ses poursuivants. Le *Ka Mate* retrace les différentes étapes de la scène et les mots employés par *Te Rauparaha* reflètent les sentiments contrastés qu'il éprouve tout au long de son épreuve :

- la descente au fond de la fosse et la crainte de l'obscurité qui préfigure la mort '*Ka mate*'
- l'espoir de la vie qui renaît en lui, quand il entend ses poursuivants s'éloigner '*Ka ora*'
- la gratitude qu'il éprouve envers son protecteur l'homme chevelu '*Tangata puruburu*', celui qui fait briller le soleil à nouveau pour lui '*Whiti te ra*'.
- la sortie de la fosse 'un pas après l'autre' '*Upane, Upane, Ka upane*'
- la remontée vers la lumière, et l'évocation réitérée de l'astre solaire, symbole de la vie '*Whiti te ra*'

Dans les lignes suivantes, on trouvera accompagnée de sa traduction, la version originale du '*Ka mate*' dans le livre '*Potatau*' du docteur Pei Te Hurinui Jones (cité par T. Karetu, 1993) :

<i>Ka mate! Ka mate!</i>	C'est la mort ! C'est la mort !
<i>Ka ora! Ka ora!</i>	C'est la vie ! C'est la vie !
<i>Tenei te tangata pururu huru</i>	C'est l'homme chevelu
<i>Nana nei te tiki mai</i>	Celui qui a fait briller
<i>I whakawhiti te ra!</i>	Le soleil à nouveau pour moi
<i>Upane! Upane!</i>	Un pas vers le haut, puis un autre
<i>Upane! Ka upane!</i>	Un pas vers le haut, puis un autre
<i>Whiti te ra!</i>	Le soleil brille !

Si l'on se fonde sur les lignes reproduites ci-dessus (en anglais le terme de *verse* est habituellement employé), le lecteur aura le sentiment (justifié) que les mots sont dénués de toute agressivité sur le plan sémantique. Toutefois, il convient de les replacer dans leur contexte original, et de préciser que ces quelques lignes font partie d'un texte dont l'esprit est foncièrement belliqueux, si ce n'est féroce comme en témoigne l'extrait suivant :

*Tirohanga nga rua rerarera* for he (the foe) will surely perish,  
*Hei a kuri ka kamukamu* and in the refuse pit will lie as food for dogs to gnaw with relish  
 (car l'ennemi périra sûrement, et sera jeté en pâture dans une fosse à ordures, où les chiens viendront le ronger et s'en repaître avec délectation).

En réalité, le *Ka mate* est une exhortation au combat où sont évoquées les vertus de bravoure et de courage des guerriers de race maorie. Dans l'étude préliminaire sur la relation voix-gestes (*op. cit.*), nous avons montré que certains mots-clés, comme par exemple '*mate*' (mort) ou '*ora*' (vie), sont ponctués de battements des mains sur les cuisses, et/ou de mouvements d'élévation des bras. Outre ces mouvements, la voix est accompagnée de diverses formes d'expression gestuelle comme la trémulation des mains '*kakapa*' ou d'expression faciale comme la dilatation des yeux '*pukana*', la fermeture des yeux '*potete*', le froncements des sourcils, ou encore la protraction de la langue '*whetero*'. De ce point de vue, on ne saurait trop insister sur la mimo-gestualité et sur l'importance du facteur corporel, et en relation avec ce qui vient d'être exposé ci-dessus, on notera que des analyses, faites au cours de campagnes électorales récentes, ont montré qu'il est plus facile d'attirer l'attention visuelle du spectateur que de retenir son attention auditive.

*Dans un discours, une personne transmet 55% du message avec le langage de son corps et les expressions faciales, 38% avec la qualité de sa voix, et seulement 7% avec le contenu du message. Autrement dit, 93% de ce qui est transmis n'a rien à voir avec ce qui est dit, et tout à voir avec l'apparence de l'orateur et sa façon de parler, (Campaigns and Election, oct.-nov. 2003, cité par Boulet-Gercourt, p. 194).*

Il est clair que dans la situation qui nous concerne, cette stratégie de communication va être appliquée *stricto sensu* par les All Blacks qui, en raison du caractère ésotérique de la langue maorie, vont accorder une importance primordiale à l'expression gestuelle en termes de mouvements de bras et de mimiques faciales, c'est-à-dire en termes d'indices visuels, sans mésestimer toutefois l'expression orale, qui va prendre la forme d'une invective où, de la combinaison de traits segmentaux et suprasegmentaux, va résulter l'émission d'un véritable cri de guerre.

### 1.3. L'exécution du *Ka mate*

Il a été mentionné plus haut que le *Ka mate* était certainement le plus connu de tous les *Haka*, étant donné qu'il est exécuté dorénavant par les All Blacks avant tous leurs matches internationaux. Toutefois, comme il est considéré, toutes proportions gardées, comme un deuxième hymne national et une marque d'appartenance à leur pays d'origine, il lui arrive d'être exécuté par des néo-zélandais (sportifs ou non), soit de leur propre initiative, soit à la demande de l'auditoire. De plus, comme le *Ka mate* se trouve être le *Haka* favori des écoles primaires et secondaires en Nouvelle-Zélande, il est certainement celui qui a été l'objet du plus grand nombre d'interprétations soit à

titre collectif, soit à titre individuel. En ce qui concerne l'interprétation qui en est faite par les All Blacks, celle-ci va être soumise à deux types d'influence.

### 1.3.1. L'influence du chef

Le chef du chœur, qui est généralement de race maorie<sup>1</sup>, mais pas forcément le capitaine de l'équipe, fait office de chef de chorale. Après l'exécution des hymnes nationaux, il rassemble ses co-équipiers autour de lui et leur prodigue ses derniers conseils. Pendant l'exécution du *Haka*, il fait face à l'équipe adverse, ses co-équipiers étant placés derrière lui, soit en arc de cercle, soit en ordre dispersé. De sa force de conviction à conduire le *Haka*, va dépendre la façon dont celui-ci va être exécuté par ses partenaires. De nombreux chefs de *Haka* se sont succédés chez les All Blacks depuis près d'un siècle, mais on s'accorde à reconnaître que c'est avec l'instauration du professionnalisme dans le rugby dans les années 1980-1990, que le *Haka* a pris une nouvelle dimension tant sur le plan de la puissance vocale du chœur que de l'expressivité gestuelle des joueurs. Parmi les chefs charismatiques qui ont joué un rôle important dans l'évolution du *Haka*, on retiendra les noms de Wayne Shelford unanimement reconnu comme *the greatest Haka leader of them all*, ainsi que de Taine Randell (T.R.), plus jeune capitaine et chef du chœur des All Blacks pendant les années 1990-2000, dont la qualité et la puissance vocale naturelle lui ont valu d'être la voix 'dominante' du *Haka* dans les enregistrements que nous avons effectués. À propos de l'histoire du *Haka*, le lecteur pourra consulter avec intérêt l'historique qui en été fait par Paul Waite sur le site <http://www.nzrugby.com/>, ainsi que celle à laquelle on peut avoir accès sur le site <http://www.rugbymuseum.co.nz/>.

### 1.3.2. L'influence de l'environnement

En plus de l'influence liée à la personnalité et au charisme du chef, il convient de mentionner l'influence de l'environnement, lequel va s'exercer consciemment ou pas sur les joueurs, au même titre d'ailleurs que sur toute personne qui se trouve en situation de communication. En effet, il a été démontré ces dernières années que le style de parole adopté par un locuteur donné, reflétait non seulement l'influence que l'environnement (physique ou humain) exerçait sur lui-même, mais également l'influence que ce dernier voulait exercer en retour sur son auditoire (*cf.* la revue critique de Ezkenazi, 1993). Dans le contexte d'environnement souvent hostile au *Haka*, où les cris poussés par la foule ont pour effet de diminuer son impact sonore, les joueurs/locuteurs vont avoir la réaction instinctive d'augmenter le volume sonore de leurs voix individuelles. Ceci explique

---

<sup>1</sup> L'actuel capitaine des All Blacks et chef du Haka, Tana Umaga, est d'origine samoane.

pourquoi le *Haka* a été quelquefois qualifié de *thundering* (tonitruant), d'*electrifying* (électrisant) pour reprendre les termes utilisés par un chroniqueur néo-zélandais bien connu (cf. *Inky's Weekly Newsletter*, July 2000) ou encore de *blood-curdling* (qui vous glace le sang) (cf. Keith Quinn, nov. 22nd 2004, *News*, That win over France !!, 2004). À la vérité, quelles que soient les conditions environnantes, un bon *Haka* ne saurait être exécuté sans que le cri soit poussé au maximum de la puissance vocale laissant le locuteur aphone, l'aphonie étant d'ailleurs considérée par Karetu comme un gage de qualité, si ce n'est d'excellence ... *if one is still able to speak after a performance, one has not given one's all* (Karetu, *op. cit.*, p. 44). Comme le maori est une langue inconnue aussi bien des adversaires des All Blacks que des spectateurs (hormis bien sûr les autochtones), il est évident que l'augmentation du volume sonore des voix individuelles ne saurait avoir une quelconque incidence sur la compréhension d'un message qui, de toute évidence, est davantage destiné à être entendu que compris. En revanche, cette augmentation du volume sonore ne va pas être sans conséquences sur la production du message verbal, et être à l'origine de nombreuses modifications qui vont altérer la structure acoustique des sons.

## 2. La réalisation phonique du Haka

### 2.1. La langue maorie

Avant d'aborder la partie segmentale, il nous semble opportun de donner ci-dessous quelques informations générales à propos des tendances qui gouvernent la langue maorie, laquelle a été étudiée de façon intensive par B. Biggs † (1921-2000). De façon succincte, nous dirons que, comparativement aux langues indo-européennes, le système phonologique du maori est simple, puisque qu'il comprend les cinq voyelles cardinales [i,e,a,o,u] et dix consonnes seulement [p,t,k,m,n,ŋ,f,r,h,w]. En revanche, sur le plan de la syntaxe et de la construction des phrases, la langue maorie est infiniment plus complexe, du fait qu'elle se distingue de la plupart des autres langues par l'absence des verbes 'être' et 'avoir'. En conséquence, les phrases vont être construites sous la forme de combinaisons de constituants marqués de façon plus ou moins formelle, d'où l'emploi de particules préposées ou postposées pour donner un sens à la phrase, utilisation intensive qui, selon Harawira (1997) n'a d'équivalent dans aucune langue du monde. En ce qui concerne la construction des phrases, les six vers du *Ka mate* illustrent cette complexité syntaxique propre à la langue maorie.

Dans les vers 1 et 2 (le second étant la répétition pure et simple du premier), le prédicat ou élément central '*Mate*' est précédé de la particule '*Ka*', dont la fonction est celle d'un marqueur de focalisation. De la combinaison de ces deux constituants résulte la formation d'une séquence

trissyllabique [*kamate*] qui signifie ‘C’est la mort’. De l’association de ce même marqueur ‘*Ka*’ avec l’autre mot-clé *Ora*, va résulter la formation d’une séquence qui est le contraire sémantique de la précédente, puisqu’elle signifie ‘C’est la vie’. Toutefois, les deux constituants *Ka Ora*, où l’on dénote la présence de trois voyelles, ne donnent pas lieu à la production d’un groupe plurisyllabique, mais au contraire à celle d’un monosyllabe [kɔ:]. Cette contraction résulte de la fusion des deux voyelles contiguës [a] et [o] en une voyelle unique d’aperture intermédiaire [ɔ], mais aussi de la chute de la deuxième voyelle finale [a]. Le vers 3 comprend deux groupes trissyllabiques ; le premier est composé de l’adjectif démonstratif ‘Tene’(i) (voici) et de l’article défini ‘Te’ (le), alors que le deuxième ne comprend que le mot lexical ‘*tangata*’ [taŋata] (homme). Dans le vers 4, après le mot ‘Nana’ [nāna] (la durée de la voyelle initiale ayant valeur phonologique en maori et impliquant la notion d’appartenance à la 3<sup>e</sup> personne du singulier), la séquence [itiki] est constituée de la particule verbale ‘i’ (qui marque le temps passé en maori) associée au verbe à l’infinitif ‘*tiki*’ (celui qui est allé chercher). Lui fait suite une autre séquence trissyllabique [mawaka] qui provient de la combinaison de la préposition ‘ma(i) (dont la dernière voyelle est occultée) et du préfixe causatif transformationnel ‘*whaka*’ (qui est la cause de quelque chose). Comme c’est essentiellement le contexte d’une phrase qui indique le temps et la condition d’une action plutôt qu’un changement de la forme verbale, la langue maorie est limitée en ce qui concerne les temps et les modes. La construction de la séquence ‘*whiti te ra*’ est l’illustration de ce qui vient d’être énoncé ci-dessus, puisqu’elle résulte de l’association d’un premier constituant verbal à l’infinitif ‘*whiti*’ (briller), d’un deuxième constituant sous forme de l’article défini ‘*Te*’ (le) et enfin du constituant nominal ‘*ra*’ dont la durée vocalique a ici valeur phonologique et signifie le soleil.

## 2.2. La transcription phonétique

En nous fondant sur nos impressions auditives, une transcription phonétique a été faite à partir de l’écoute du signal de parole copié d’après la bande son de la cassette vidéo. Le *Haka* peut être subdivisé en deux parties. La première partie est une harangue prononcée par le chef T. Randell, le chœur n’intervenant que de façon secondaire dans la partie terminale, alors que la deuxième partie concerne le *Ka mate* proprement dit, repris en chœur par les vingt-deux joueurs (les quinze titulaires du début de match et les sept remplaçants).

Selon les versions que nous avons trouvées dans les livres ou dans les articles de la presse sportive, la longueur de la harangue est variable. Dans le livre de T. Karetu (*op. cit.*), la version est longue puisqu’on y fait référence à plusieurs membres ou parties du corps comme les mains, les cuisses, la poitrine, les genoux, les hanches et les pieds. En revanche, dans la version reproduite par Harawira (1994) dans son manuel d’initiation à la langue maorie, celle-ci est brève puisque ne sont évoqués

que les mains ‘ringaringa’ et les pieds ‘waewae’. Si l’on se fonde sur la transcription phonétique présentée ci-dessous, il ressort que c’est cette dernière version (écourtée et édulcorée) qui a été interprétée par le chef.

<i>Version écrite longue (Karetu, 1994)</i>		<i>Version écrite brève (Harawira, 1994)</i>
<i>Ringa pakia</i>	(frappez des mains sur les cuisses)	<i>Ringa pakia</i>
<i>Uma tiraba</i>	(gonflez la poitrine)	<i>Waewae takabia kia rite</i>
<i>Turi whatia</i>	(pliez les genoux)	<i>E kino nei boki</i>
<i>Hope whai ake</i>	(déhancez-vous)	
<i>Waewae takabia kia kino</i> (frappez du pied aussi fort que vous le pouvez)		

<i>Transcription phonétique de la harangue</i>	
Le chef	[karite # karite # ?ikamau]
Le chœur	[?u:] #
Le chef	[?arĩarĩapakia#majwajtrijakakino # ?aki:]
Le chœur	[?akinonehoki]

<i>Version écrite (Dr Pei Te Hurinui Jones) Transcription phonétique</i>	
<i>Ka mate! Ka mate! Ka ora! Ka ora!</i>	[?a:kamatekamateko:]
<i>Ka mate! Ka mate! Ka ora! Ka ora!</i>	[?a:kamatekamateko:]
<i>Tenei te tangata puburuhuru</i>	[?a:tejnetetāŋata # u:]
<i>Nana nei te tiki mai I whakawhiti te ra</i>	[na:naitikimawakafititea:]
<i>Upane! Upane!</i>	[?aupā # ?aupā]
<i>Upane! Kaupane! Whiti te ra!</i>	[?aupānekupānefititea: # hi:]

### 2.3. La distribution et la symbolisation des sons

Au cours des années 1950-1980, plusieurs linguistes, versés dans l’art de la phono-stylistique, se sont attachés à montrer qu’il existait sur le plan du symbolisme phonétique une relation entre les émotions éprouvées par les écrivains ou poètes et la nature des segments phoniques qu’ils utilisaient dans leurs écrits pour exprimer leurs émotions. Dans la section 1.3.2., nous avons signalé que l’influence du *Haka* s’exerçait sur l’adversaire à la fois par le geste et par la voix. En ce qui concerne cette dernière forme d’influence vocale, celle-ci ne peut être pleinement effective que si le message est transmis de la voix la plus forte possible. Comme il a été établi que la majeure partie de l’énergie du signal vocal se trouve contenue dans les segments vocaliques, il est légitime de penser que ces derniers ont été privilégiés pour produire l’effet auditif et psychologique recherché.

Si cette prédominance des segments vocaliques s'exprime en termes de pourcentages d'occurrence (61%) par rapport à leurs homologues consonantiques (39%), nous verrons dans le chapitre consacré à l'analyse acoustique que cette prédominance apparaît encore plus nettement en termes de rapports de durée entre les segments.

### 2.3.1. Les voyelles

Le tableau 1 illustre la fréquence d'occurrence des voyelles produites par le chef durant la harangue, par les vingt-deux joueurs pendant le *Ka mate*, ainsi que durant tout le *Haka*. L'examen des données quantitatives brutes donne lieu aux remarques suivantes :

**Tableau 1**  
*Fréquence d'occurrence des voyelles (en %)*

	Harangue	Ka Mate	Haka
/i/	33.33	20.44	26.88
/ɛ/	9.09	15.90	12.49
/a/	45.45	48.88	47.16
/ɔ/	9.09	5.79	7.44
/u/	3.03	9.03	6.03

La remarque principale porte sur la prééminence de la voyelle [a] dont le nombre d'occurrences est proche de 50%, ce qui signifie que dans un cas sur deux, c'est la voyelle [a] qui a été produite, le mot *Haka* étant d'ailleurs la parfaite illustration de cette prééminence. Dans le contexte d'environnement bruité que nous avons décrit, cette prédilection s'explique en fonction des propriétés acoustiques intrinsèques de la voyelle.

En raison d'une durée intrinsèque qui est supérieure à celle de toutes les autres voyelles, elle se prête à être produite sous la forme d'une voyelle de (très) longue tenue, ce qui va naturellement favoriser sa perception. De plus, comme son intensité spécifique se trouve être la plus élevée de toutes les voyelles, elle va être la plus audible surtout dans des conditions de bruit. Enfin, comme un tractus vocal partiellement fermé ne saurait favoriser la propagation d'une onde sonore dans son amplitude maximale, il importe que le tractus soit le plus ouvert possible, configuration articulaire qui est précisément celle de la voyelle [a]. Sur le plan acoustique, toutes les conditions sont donc réunies pour que cette voyelle se trouve privilégiée sur le plan distributionnel.

Notre opinion se trouve confortée par celle de nombreux auteurs qui ont mis en avant sa sonorité ainsi que sa force vocale pour exprimer des sentiments violents comme la haine ou la colère. Dans le contexte d'affrontement physique que préfigure le *Haka*, sa production est propice à donner un caractère dur, masculin et viril aux syllabes qui la contiennent. La fréquence d'occurrence de cette voyelle est due au fait qu'elle se rencontre TOUJOURS à l'attaque des six vers, mais aussi quelquefois en position finale (vers 4, 5). Dans cette dernière position, il arrive qu'on trouve aussi la voyelle ouverte [ɔ] dont l'intensité spécifique, quoique inférieure à celle de [a], est relativement élevée (Rossi, 1978). En ce qui concerne les graphèmes E et O, leur production phonique est plus proche de l'allophone ouvert que de l'allophone fermé. C'est pourquoi ces voyelles sont proches dans la langue maorie de celles des mots 'bed' [bɛd] ou 'porc' [pɔ:k] de l'anglais britannique (Harawira, *op. cit.*). Dans le *Haka*, ce sont des allophones ouverts qui sont perçus en position finale, notamment la voyelle finale [ɛ] dans la harangue [karitɛ] ou encore la voyelle [ɔ] à la fin des deux premiers vers dans le monosyllabe [kɔ:] suite à l'amuïssement de la syllabe [ra] (*cf.* remarque précédente). Cette assertion confirme la justesse des observations de Straka (1964) qui a remarqué d'après l'examen de radiographies et de palatogrammes, que l'augmentation de l'énergie articulatoire avait un effet inverse sur les voyelles et les consonnes, en ce sens que les premières témoignaient d'une tendance à l'ouverture, alors que les secondes s'en distinguaient par une tendance à la fermeture. Dans le chapitre consacré à l'analyse acoustique (section 4.6.2), nous verrons que cette tendance à l'ouverture est la règle, puisque ce sont des voyelles ouvertes qui sont généralement produites (et perçues) en position finale de mot notamment dans le vers 3 [tejnetɛ], ainsi que dans le vers 6 [ʔaupaneɪkupane].

Il résulte de l'effet combiné de la prééminence de [a] et de la tendance à l'ouverture de [ɛ] et de [ɔ], qu'une majorité de voyelles ouvertes (70.57%) se dégage par rapport aux voyelles fermées (29.43%). Parmi ces dernières, la voyelle [i], *la plus faible et la moins sonore des voyelles*, (Fónagy, 1983, p. 76) est relativement fréquente, quoique sa fréquence d'occurrence (20.44%) se situe nettement en retrait par rapport à la voyelle [a] (48.88%). Malgré une sonorité moindre, elle occupe une place relativement importante dans les voyelles du *Haka*, parce que son timbre aigu évoque le concept de 'lumière', de 'clarté' ou de 'brillance', *cf.* la séquence '*whiti te ra* [fititea] (le soleil qui brille) aux vers 4 et 6, ainsi que la séquence [itiki] (vers 4).

L'autre voyelle fermée [u] trouve également sa place dans ce contexte d'antagonisme, et le cri poussé par les ABs dans la harangue préfigure le combat qui va s'ensuivre. En effet, c'est la voyelle dont le timbre grave et sombre rappelle *la voix des mâles robustes...*, (Fónagy, *op. cit.*, p. 78). À la différence de la voyelle [i] qui est brève, la voyelle [u] (vers 3) et le cri (de timbre grave) dans la harangue sont produits sous la forme de voyelles tenues. Se référant à la langue hongroise, Fónagy



note une tendance à l'allongement et au renforcement des voyelles sous l'effet d'un sentiment ou d'une émotion forte. La pertinence de cette remarque est vérifiée dans le *Haka*, où l'allongement se manifeste sur l'axe syntagmatique non seulement par une durée inhabituelle des voyelles, mais aussi par une restructuration complète de la séquence 'puburuburu' produite sous la forme de deux voyelles séparées par une pause [hu:# hu:] (vers 3). Dans le même sens, Jespersen (1922) avait noté une tendance naturelle à l'allongement ou au renforcement de certains sons sous l'influence d'un sentiment très fort, afin d'intensifier l'effet du mot prononcé. Sa remarque ... *il est même possible que dans certains cas, un allongement vocalique dû à un facteur psychologique modifie complètement le mot* (Jespersen, *op. cit.*, p. 390) trouve ici sa pleine justification, si l'on se réfère à la déformation de la séquence indiquée ci-dessus.

Si cet allongement des voyelles est donc associé à l'intensification d'une action ou à l'expression d'une émotion violente, un autre élément de dureté est introduit par le biais de voyelles enchaînées, dont les *chocs vocaliques rendent les vers merveilleusement rudes* (Morier, 1961, p. 481). Fónagy utilise pour sa part la métaphore de voyelles qui *s'entrechoquent et dont la production est la plupart du temps accompagnée d'un coup de glotte*, (Fónagy, *op. cit.*, p. 95). À vrai dire, cette impression auditive n'est pas nouvelle, et a été ressentie de longue date, plus précisément dès l'antiquité romaine, puisque Quintilien évoquait *la collision de voyelles qui se heurtent et produisent un vide, un arrêt qui fait que la parole se traîne péniblement*, (De Institutione Oratoria, IX, p. 4, cité par Fónagy, *op. cit.*, p. 98).

En réalité, la présence de voyelles enchaînées, appelées également voyelles en hiatus, répond à une norme et peut être considérée comme l'un des traits dominants du système vocalique de la langue maorie. Toutefois, cette dernière se différencie de la plupart des langues anglo-saxonnes, parce que chacune des deux voyelles conserve son timbre et la deuxième n'est généralement pas diphtonguée, ... *When two vowels come together in a word, each must be given its own sound*, (Harawira, *op. cit.*, p. 9). De même, Mc Lagan *et al.* (2004a) font remarquer dans une étude récente sur la prononciation du maori contemporain que, en raison de contraintes phonotactiques, de nombreuses voyelles se suivent dans cette langue sans former pour cela des diphtongues phonologiques. Dans le *Haka*, les voyelles enchaînées sont nombreuses, n'en serait-ce pour preuve l'association de la voyelle [a] avec plusieurs voyelles de timbre différent, par exemple avec la voyelle [i] dans [naitiki] ou [ia] dans [pakia] ou encore avec la voyelle [u] dans [ʔikamau] et [ʔaupā]. On notera également la rencontre des voyelles [ea] dans [fititea] suite à l'élision de la consonne à battement [ɾ]. Outre ces voyelles enchaînées, le *Haka* se distingue également par la récurrence de voyelles de même timbre, lesquelles ne sont séparées les unes des autres que par de brèves ruptures consonantiques, par exemple le long continuum vocalique formé de l'association des trois voyelles [a-a-a] dans [ʔakama] [tāŋata] ; [ɛ-ɛ-ɛ] dans [tejnetɛ] ; [i-i-i] dans [itiki] ou

encore des deux voyelles [i-i] dans [fiti] et [u-u], ces deux dernières voyelles étant entrecoupées d'une pause virtuelle (*cf.* section 5.6.2.2.). Cette association de voyelles de même timbre implique que les changements de couleur vocalique sont relativement rares dans la harangue où, à l'exception de la voyelle [ɔ], ne se rencontrent que les seules voyelles [a] et [i] dans la séquence [ʔarĩŋarĩŋapakia#majwajtri # ʔakakino # ʔaki:]. Cette alternance de voyelles [i-a] se retrouve pareillement dans le vers 4 du *Ka mate* [na:naitikimawakafitea:], l'exception à la règle étant dans ce cas la voyelle [e], alternance que l'on retrouve dans le vers 5, et qui est due à la réitération de la séquence [ʔaupa # ʔaupa].

### 2.3.2. Les consonnes

**Tableau 2**  
*Fréquence d'occurrence des consonnes (en %)*

		Harangue	Ka Mate	Haka
occlusives	/p/	3.57	9.76	7.35
sourdes	/t/	10.77	31.71	23.53
	/k/	32.14	21.95	26.47
<b>Total</b>		<b>46.48</b>	<b>63.42</b>	<b>57.3</b>
nasales	/m/	7.14	12.20	10.29
	/n/	10.77	12.20	10.29
	/ŋ/	7.14	2.44	4.41
<b>Total</b>		<b>25.05</b>	<b>26.84</b>	<b>24.99</b>
constrictives	/f/		4.76	3.03
<b>Total</b>				<b>3.03</b>
glides	/j/	7.14	2.38	4.41
	/w/	3.57	2.38	2.94
	/r/	17.86		7.35
<b>Total</b>		<b>28.57</b>	<b>4.88</b>	<b>14.70</b>

Si la prédominance de la classe des voyelles ouvertes semble solidement établie, il en est de même de la classe des consonnes occlusives sourdes, comme en attestent les pourcentages d'occurrence recensés ci-dessous (tableau 2). Toutefois, on se doit de reconnaître qu'il existe une relation entre la distribution des consonnes dans le *Haka* et la fréquence d'usage des phonèmes qui en est faite

en fonction du système phonologique de la langue maorie. Le fait que les consonnes occlusives sourdes /p-t-k/ ainsi que les consonnes nasales /m-n-ŋ/ soient nettement majoritaires dans le *Haka*, s'explique du fait que les trois consonnes labiale, dento-alvéolaire et vélaire se retrouvent toutes dans le système consonantique du maori, alors que les consonnes sonores en sont totalement absentes et qu'il n'existe qu'une seule consonne constrictive [f].

Au même titre que les voyelles ouvertes, la prédominance des consonnes occlusives (53%) est susceptible de s'expliquer en relation avec les velléités agressives qui sont exprimées dans le *Haka*. En effet, si l'on se fonde sur les résultats de tests sémantiques portant sur l'association métaphorique d'un son avec une émotion donnée, on constate que les occlusives sourdes /p,t,k/ se situent en haut de l'échelle de dureté, mais inversement en bas de l'échelle d'euphonie des consonnes. Sur ce point, les psycho-stylisticiens s'accordent à reconnaître que les occlusives sourdes, en particulier l'occlusive vélaire /k/, sont associées à la *pulsion agressive* (Fónagy, *op. cit.*, p. 88), à la *dureté* et à la *masculinité* (selon l'opinion exprimée par le grammairien anglais Thomas Smith), et enfin que ces consonnes sont jugées comme *dures*, *malsonnantes*, *cacophoniques à l'oreille* (Grammont, 1933). Ces auteurs insistent sur le bruit *sec* et *éclatant* provoqué par l'explosion de la consonne occlusive au contact de la voyelle /a/, en particulier quand celle-ci se trouve en position initiale absolue ou en contexte de groupe consonantique (occlusive + latérale ou vibrante). En effet, cette consonne se rencontre fréquemment à l'initiale de mots qui évoquent le bruit fort de quelque chose qui résonne, qui se casse ou qui vole en éclats, par exemple les mots 'cataracte, cascade, clameur, fracas' en français, les mots 'crash, clash' en anglais ou encore le mot 'Klang' en allemand.

En ce qui concerne la langue maorie et le *Haka*, l'attention du lecteur ou de l'auditeur ne peut être qu'attirée par la récurrence systématique du phonème /k/ au début du *Ka mate*. En effet, le premier constate *de visu* la succession des quatre syllabes /ka/ en début de texte, alors que l'auditeur perçoit le son /k/ à six reprises au cours des deux premiers vers. Cette succession de consonnes 'dures' associées à l'expression de sentiments violents comme la colère ou la rage est fréquente en langage poétique. Un exemple nous en est donné dans la pièce de théâtre *la Cantatrice chauve* de Ionesco, où la répétition obsessionnelle du son /k/ dans les mots 'cascade, cacade, cacatoès' contribue à renforcer le sentiment de haine entre les deux protagonistes sur scène. On signalera à ce propos qu'un procédé symbolique similaire est utilisé par P. Verlaine dans le poème *Les Injures*, où selon l'expression de Fónagy (*op. cit.*, p. 66) '*le vers est saccagé, la phrase déchirée et mise en morceaux par le nombre d'occlusives sourdes*'.

Dans la situation présente, la valeur impressive de la consonne /k/ (et par extension celle de la syllabe /ka/) se trouve renforcée du fait que sa production est synchronisée avec les battements provoqués par le contact des mains sur les cuisses. De ce contact résulte un claquement sec qui a

pour effet de renforcer l'impact sonore de la syllabe tout entière. Cette remarque s'applique également à l'occlusive sourde /p/ dans la séquence [ʔau'pa] où la consonne est produite en synchronisation avec les bruits résultant du double contact de la paume de la main sur l'avant-bras opposé. En fait, les deux claquements sont synchronisés avec la production de la syllabe /pa/ et contribuent à la mettre en relief sur le plan perceptif.

Quoique le pourcentage d'occurrence de /t/ (31.71%) soit nettement supérieur à celui de /p/ (9.76%) et de /k/ (21.95%), cette consonne ne bénéficie pas d'un relief sonore égal à celui de ses homologues sourdes. En effet, si les syllabes /tɛ/ dans [tejnɛtɛ] et /tã/ dans [tãŋata] sont toutes deux perçues comme accentuées en raison de leur position initiale forte, en revanche elles sont moins saillantes sur le plan perceptif, car leur production n'est pas ponctuée d'un claquement sec comme c'est le cas de [k] ou de [p].

Un autre son qui a également reçu le qualificatif de consonne 'dure' est le /r/ apical, associé pour sa part aux concepts de rugosité, virilité ou masculinité, « *le r prend tout son relief lorsqu'il s'agit d'actions violentes et véhémentes* » (*quandoque exponuntur res acerbae ac vehementes*) (Vossius, cité par Fónagy, *op. cit.*, p. 96). Cette opinion est partagée par le poète hongrois S. Petöfi, qui, outre les consonnes dites dures comme /p-t-k/, a privilégié l'utilisation de la consonne /r/ (dont le roulement est interprété par l'auteur comme une menace) pour exprimer la haine des patriotes hongrois envers l'occupant ottoman. En ce qui concerne la langue maorie, c'est la variante apicale à un battement (flap) (considérée comme plus expressive que la variante uvulaire) qui est la plus fréquente, quoique la variante à plusieurs battements (trill) se rencontre également dans certains dialectes de l'île du sud (Biggs, 1971). Toutefois, si on retrouve le graphème 'R' à neuf reprises dans le texte, en revanche il n'est produit (et perçu) que de façon occasionnelle. À notre opinion, cette divergence entre l'écrit et l'oral s'explique en fonction de deux critères. Le premier est d'ordre environnemental, et est lié au fait que si la connotation de 'rugosité' du son [r] peut effectivement être mise en valeur dans la lecture d'un poème dans le silence feutré d'un salon, il en va tout autrement dans le *Haka*, où ladite connotation est caduque en raison du vacarme assourdissant qui règne dans un stade. Le deuxième critère a trait aux contraintes physiologiques (et sociolinguistiques) qui font que la réduction du geste alvéolaire, et *a fortiori* l'élision complète du son [r] sont fréquentes, quand le locuteur change de style de parole et passe d'une parole soignée à une parole familière (Labov, 1964) ; Browman and Goldstein (1990))<sup>2</sup>.

---

<sup>2</sup> Le son [r] semble avoir connu une période de défaveur en France à l'époque du Directoire, puisque étaient privilégiés dans les salons parisiens les 'inc-oyables' et les 'me-veilleux'.

À l'instar des consonnes occlusives, les seules consonnes que l'on retrouve dans chaque vers du *Ka mate* sont les consonnes nasales, d'où un pourcentage d'occurrence élevé (25%). Selon une vieille tradition rhétorique et poétique, ces consonnes étaient associées à la douceur et à la suavité, d'où le qualificatif de 'phonèmes doucereux' et une utilisation intensive dans les poèmes tendres ou érotiques (Fónagy, *op. cit.*, p. 75). Dans le *Haka* qui n'entre pas, loin s'en faut, dans ce genre de littérature, leur présence répond à un autre besoin. En fonction de leur sonorité, elles se lient facilement aux segments vocaliques contigus avec lesquels elles se trouvent en état de coarticulation avancée ; d'autant plus que, se trouvant le plus souvent en position interne de mot, elles participent à la fluidité du continuum après la syllabe accentuée, et facilitent l'enchaînement des syllabes en position post-accentuelle, par exemple les séquences trisyllabiques '*Kamate*', '*Teneite*', '*Kaupane*' etc.

## 2.4. La relation graphie-phonie

Les différences majeures que nous avons constatées entre les versions écrite et orale portent essentiellement sur l'adjonction de voyelles et sur l'élision de syllabes. Toutefois, avant d'examiner en détail ces deux points, il convient de remarquer que la prédominance des voyelles ouvertes et des consonnes sourdes qui contribuait à renforcer le caractère de violence et de dureté du *Haka*, est essentiellement due à la répétition de quelques mots-clés qui reviennent comme une litanie.

Dans le *Ka mate*, les mots que l'on retrouve le plus souvent sont '*mate*' (x4), '*ora*' (x2), '*upane*' (x4), '*whiti te ra*' (x2). Cette répétition n'est pas fortuite et retrace la chronologie des événements qui voient *Te Raupara* passer alternativement de la crainte de la mort à l'espoir de la vie. En effet, nous avons vu que le premier mot du texte évoque sur le plan sémantique le concept de la mort '*mate*', alors que le deuxième '*ora*' évoque son contraire, c'est-à-dire la vie. Cette dualité de la mort et de la vie est constamment présente dans le *Ka mate*, puisque celui-ci commence par le mot '*mate*', et se termine par la séquence '*whiti te ra*' qui évoque la brillance du soleil, symbole de la vie. La répétition des mots '*upa*' et '*upane*' n'est pas fortuite et implique une notion de continuité, puisqu'ils évoquent la remontée et la sortie 'un pas après l'autre' de *Te Raupara* hors de la fosse.

En ce qui concerne la harangue, c'est la version brève qui a été interprétée par le chef. Toutefois, elle diffère sensiblement du texte du fait que sont amalgamés (quelquefois dans le désordre) les mots ou séquences de mots que l'on retrouve chez Harawira (*op. cit.*) (section 2.2). C'est ainsi que la harangue débute par la répétition de *Ka rite* [*karite*] suivi de la séquence [*?ikamau*] qui n'est pas mentionnée dans le texte. Ensuite, la harangue est plus ou moins conforme à la version écrite (à l'exception de '*takabia*') et se termine par la séquence [*?akakino # ?aki*] reprise en chœur par le groupe. En fait, la harangue constitue moins un soliloque de la part du chef, qu'un dialogue entre

ce dernier et ses équipiers qui lui répondent de deux façons. La première réponse prend la forme d'une voyelle tenue [hu:] ; elle évoque sur le plan perceptif le hululement de la chouette ou le *rugissement des fauves féroces* (Peterfalvi, 1970) et fait écho à la voyelle finale de la séquence [ʔikamau] produite par le chef. La deuxième réponse a la forme d'une séquence de plusieurs syllabes [ʔakinonehoki:] qui fait écho à la même séquence produite par le chef ; toutefois, elle en diffère du fait que l'on perçoit une légère variation de timbre entre la production du chef [haki:] et la production du chœur [hoki:].

#### 2.4.1. L'adjonction de voyelles

Sur le plan étymologique, le mot '*Haka*' résulte de l'association de deux syllabes, la première '*Ha*' ayant trait au concept de 'respiration' alors que la deuxième '*Ka*' implique pour sa part le 'déclenchement' d'une action. Ainsi le mot *Haka*, en plus du concept de danse auquel il est généralement associé, évoque le 'déclenchement de l'acte respiratoire'. L'observation des films vidéo confirme cette interprétation ; en effet, on constate que, conformément à l'une des recommandations sous la forme de la séquence '*Uma Tiraha*' (gonflez la poitrine), les joueurs se livrent à des exercices respiratoires en termes d'inspirations et d'expirations. À ce sujet, il est permis de s'interroger sur les raisons de ces exercices respiratoires. À notre opinion, ils ont pour but soit d'insuffler de l'énergie au corps tout entier, soit d'évacuer la tension psychologique par la régulation du rythme cardiaque, soit enfin d'accumuler un maximum d'air phonatoire dans les poumons avant de pousser le cri de guerre. Dans cette dernière optique, il est évident que ce cri ne peut être puissant que si l'entrée d'air phonatoire a été favorisée par une large ouverture de l'orifice buccal. L'examen de l'oscillogramme synchronisé avec l'image, montre que, à l'instant de la première impulsion glottique, c'est-à-dire au moment même où la glotte se ferme et où les cordes vocales entrent en vibration, l'ouverture buccale est effectivement maximale. Or, c'est précisément la voyelle [ʔa] qui est produite à l'attaque des différents groupes de souffle de la harangue [karite], [ʔarīŋa], [ʔakinone]. En ce qui concerne cette dernière séquence, on notera la divergence entre l'oral et l'écrit, du fait que la production du phonème [a] n'est pas en conformité avec la présence du graphème E dans la version d'Harawira (*op. cit.*). La production d'une voyelle ouverte à l'attaque de chaque vers du *Ka mate* est partout la règle, à l'exception toutefois du vers 4 qui commence par la consonne nasale [n], cf. le mot '*nana*'. Cependant, il convient de noter que, en raison de la coarticulation engendrée par l'anticipation vocalique et des différences intrinsèques de durée et d'intensité entre les deux sons, c'est la voyelle [a] qui est saillante sur le plan perceptif et non pas la consonne nasale.

Nous avons signalé plus haut, que cette même voyelle [a] se retrouve à la finale des vers. Ce cas de figure se rencontre au vers 4 qui se termine par deux voyelles en hiatus dans [fititea] ainsi qu’au vers 5, cf. la syllabe [pa] dans [ʔaupā]. Quand ce n’est pas la voyelle [a], c’est la voyelle la plus proche [ɔ] qui est produite suite à l’élision de la syllabe finale [ra] (vers 1 et 2), ce qui tend à confirmer la tendance à l’ouverture des voyelles dites intermédiaires, (Harawira, *op. cit.*). À la fin du vers 3, on rencontre une voyelle tenue fermée [u:]. Enfin, le cri qui conclut le vers 6 est synchronisé avec le saut ponctué par une élévation des bras vers le ciel. Dans plusieurs versions écrites du *Haka*, ce dernier son est transcrit sous la forme du graphème I. Quoique le timbre de ce cri collectif soit difficile à déterminer sur le plan perceptif, il sera établi sur le plan acoustique que sa structure de formants l’apparente effectivement à une voyelle antérieure [i] de longue durée.

#### 2.4.2. L’élision de syllabes

Au même titre que l’adjonction de voyelles, l’élision de syllabes joue un rôle important dans l’opposition graphie-phonie. Le premier cas d’élision se rencontre dans la harangue, où d’après la version écrite, le mot bi-syllabique *waewae* (pied) est suivi de la séquence ‘*takabia kia rite*’. Or, si l’on se fonde à la fois sur l’audition du signal sonore et sur l’analyse acoustique, on constate que cette dernière séquence n’a pas été perçue et produite, en tant qu’une séquence polysyllabique, mais en tant qu’un monosyllabe que nous avons transcrit [tri:]. Il semble donc que nous ayons affaire à un phénomène d’élision de plusieurs syllabes, où pas moins de cinq à six d’entre elles se trouvent condensées en une syllabe unique. Un autre cas d’élision de syllabes se présente au vers 3 avec la séquence ‘*puburuburu*’. Comme la précédente, cette séquence pluri-syllabique va se trouver condensée en une voyelle unique [hu:] qui sera néanmoins perçue en deux segments lesquels sont séparés par une pause virtuelle (cf. remarque précédente).

Plus fréquente est l’élision de syllabes uniques. Un exemple nous en est donné dès la fin du premier vers du *Ka mate* où la chute de la syllabe finale [r] entraîne la fusion de la voyelle [a] du marqueur de focalisation ‘*Ka*’ avec la voyelle [o] du mot suivant ‘*Ora*’. La fusion de ces deux voyelles va donner lieu à la production d’une voyelle intermédiaire sur le plan de l’aperture entre [o] et [a], c’est-à-dire [ɔ:]. Concernant la chute de la syllabe [ra], on remarquera que ce phénomène de dévoisement de la syllabe finale est courant en maori, et qu’il est généralement accompagné d’une baisse concomitante de l’intensité et de la fréquence fondamentale (Biggs, 1971). Un autre cas d’élision apparaît au vers 4 avec l’élision des syllabes simples ‘*nei*’ ou ‘*Te*’, de [i] dans ‘*mai*’ ou encore ‘*T*’. Il en résulte une homogénéité de couleur vocalique puisque la séquence est constituée d’une alternance des deux voyelles [a] et [i], cf. [na: na itiki mawaka fiti..a]. Cette combinaison harmonieuse de voyelles (due ici à l’élision des monosyllabes ‘*nei*’ et ‘*Te*’), se retrouve tout au long

du *Haka*, et explique cette impression d'euphonie et de musicalité que l'on éprouve à l'écoute de la langue maorie, '*Maori pay much attention to euphony, and will therefore employ the most melodious-sounding word rather than follow a particular rule* (Harawira, *op. cit.*, p. 55). Contrairement à la représentation graphique, le vers 5 est marqué par l'élision de la syllabe finale '*ne*' dans le mot '*Upane*', l'élision de la syllabe terminale étant compensée par l'émission d'une voyelle initiale [a] et donnant lieu à la formation de voyelles en hiatus, cf. les deux séquences [ʔaupa # ʔaupa] qui sont entrecoupées d'une pause. En revanche, la syllabe écrite '*ne*' est produite dans le même mot au vers 6 ; dans ce cas, les séquences [ʔaupaneʔupane] sont enchaînées et ne sont séparées par aucune pause comme c'est le cas dans le vers précédent.

En résumé, nous dirons qu'il existe de nombreux cas de divergence entre la production orale des All Blacks et les différents textes écrits du *Haka* que nous avons rencontrés. Dans le chapitre suivant, nous allons essayer de confirmer la justesse de nos impressions subjectives, en nous fondant sur le recueil de données quantitatives grâce à une analyse acoustique fine du signal de parole.

### 3. L'analyse acoustique

#### 3.1. Le protocole expérimental

Les études préliminaires nous avaient permis de montrer qu'il existait une corrélation étroite entre les deux composantes orale et gestuelle. Quoique la présente étude soit consacrée essentiellement à la première de ces deux composantes, il va de soi que l'une ne peut se concevoir sans l'autre, et qu'en de nombreuses occasions, nous serons amenés, si nous ne l'avons déjà fait, à nous référer au geste sous-jacent qui accompagne la voix. C'est pourquoi il nous a semblé utile de revenir dans cette section sur les difficultés que nous avons rencontrées pour l'analyse synchrone de ces deux composantes et l'établissement d'une relation satisfaisante entre l'image et le son.

Notre première tâche (la plus agréable et la plus facile !!!) a porté sur l'enregistrement sur cassettes vidéo de plusieurs rencontres opposant les All Blacks à différentes équipes, rencontres diffusées sur la chaîne de télévision cryptée Canal Plus. Afin de satisfaire aux conditions d'analyse, les enregistrements devaient répondre à deux critères de qualité visuelle et auditive.

##### 3.1.1. La qualité visuelle de l'image

Pour la retransmission télévisée de rencontres de rugby ou de tout évènement sportif en général, le matériel technique dont dispose le réalisateur comprend plusieurs caméras fixes ou mobiles placées à l'intérieur du stade. Parmi les images qui sont transmises simultanément par les 'cameramen', le réalisateur va choisir (en studio ou dans le car de régie) celle qui est la plus appropriée pour rendre compte de l'action à un moment donné. Pendant l'exécution du *Haka*, plusieurs options lui sont



offertes. Il peut opter soit pour un plan ‘individuel’ sur un joueur dont le comportement gestuel ou la mimique faciale lui semblent particulièrement télégraphiques, soit pour un plan ‘collectif’ sur un groupe de joueurs ou sur l’équipe tout entière, soit enfin pour un plan qui illustre le comportement et les réactions de l’adversaire. En règle générale, le réalisateur privilégie l’aspect dynamique à l’aspect statique, c’est-à-dire qu’il alterne les angles de prises de vue sur les joueurs, *cf.* vue de face, vue de profil, quelquefois vue de dos, afin de transmettre un maximum d’informations visuelles au téléspectateur. De ce fait, le réalisateur ne s’attarde que rarement sur le même joueur ou sur le même groupe de joueurs ; en d’autres termes, l’image est la plupart du temps variable. Cette variété d’images, si elle est appréciée du téléspectateur qui assiste à l’exécution du *Haka* grâce à divers angles de vue, ne l’est pas forcément du chercheur dans la mesure où le signal de parole émis par les joueurs ne correspond pas nécessairement à l’image projetée sur l’écran, ce qui entraîne forcément une rupture de correspondance entre le son et l’image. De plus, quand les deux se trouvent être synchronisés, cette synchronisation est la plupart du temps brève. Dans ces conditions, il nous a fallu enregistrer plusieurs rencontres et visualiser de nombreuses séquences avant d’associer un comportement gestuel donné à la production d’une unité phonique, d’une syllabe ou d’une séquence de syllabes spécifique.

### 3.1.2. La qualité auditive du son

L’autre difficulté majeure à laquelle nous avons été confrontés concerne la qualité auditive d’un signal de parole enregistré en champ libre. Comme ce signal est souvent perturbé par des interférences acoustiques diverses (cris ou sifflets de la foule, battements des mains, martèlement des pieds etc.), il est difficilement audible et la plupart du temps impropre à l’analyse acoustique. Dans ces conditions d’environnement sonore défavorable, nous avons privilégié un enregistrement effectué en Nouvelle-Zélande où le *Haka* est évidemment respecté, et même quelquefois exécuté dans un silence quasi religieux. Un autre facteur de qualité de l’enregistrement réside dans le nombre et l’emplacement des micros. Ces deux facteurs favorables, d’une part le silence (relatif) de la foule et d’autre part la disposition adéquate des micros autour du chœur, se sont trouvés réunis à l’occasion d’un match opposant les All Blacks à l’Écosse, à Wellington en juillet 2000. La harangue était enregistrée par l’intermédiaire d’un micro maintenu près de la bouche du chef par un preneur de son, alors que le *Ka mate* était enregistré grâce à plusieurs micros placés sur des supports fixes à proximité immédiate des joueurs. Compte tenu de ces conditions favorables, c’est la bande-son originale de la cassette vidéo qui a été retenue. En un deuxième temps, celle-ci a été copiée sur une bande magnétique, et a constitué le document sonore de base utilisé à la fois pour la transcription phonétique, l’analyse acoustique et enfin l’étude perceptive des faits prosodiques.

### 3.2. Le filtrage et le traitement du signal

Quoiqu'il ait été enregistré dans des conditions plutôt favorables, le signal de parole n'en était pas pour autant exempt de toute interférence acoustique. En effet, une foule de plusieurs dizaines de milliers de spectateurs ne saurait être jamais totalement silencieuse et une analyse effectuée sur une portion de signal d'une durée de 100ms préalablement à l'émission du signal de parole, a révélé la présence d'un bruit de fond de l'ordre de 20dB, bruit particulièrement intense dans la gamme fréquentielle 800-1200 Hz. Grâce au logiciel *Sound Forge*, ce bruit a été réduit de 10dB, réduction qui a eu pour effet d'améliorer la qualité auditive du signal. L'analyse acoustique a été fondée sur l'utilisation du logiciel *Phonedit*. En un premier temps, le signal analogique a été numérisé, segmenté, et étiqueté par l'auteur, par exemple RIN (1<sup>e</sup> partie de la harangue), KA1 (premier vers du *Ka mate*), KA2 (deuxième vers), TE3 (troisième vers) etc. En un deuxième temps, des spectrogrammes numériques en bande large (300 Hz) ont été faits sur une gamme fréquentielle 0-4.0 Khz, laquelle s'est avérée suffisante du fait que les consonnes du *Haka* ne comprenaient qu'une seule consonne fricative<sup>3</sup>. La mesure des fréquences des formants des voyelles a été faite par l'application de la Transformée de Fourier rapide avec une fenêtre de *Hamming* de 32 ms à 512 points. Les spectrogrammes numériques ont servi à la mesure de la durée des vers, des groupes rythmiques, des syllabes, des pauses, ainsi que des tenues vocaliques et consonantiques. L'analyse des propriétés prosodiques a été fondée sur l'utilisation du logiciel *Praat* pour la visualisation des configurations de fréquence fondamentale (F0) et d'intensité (A0), ainsi que pour la mesure des variations de ces paramètres qui participent de ces configurations.

## 4. Les propriétés segmentales des sons du Haka

### 4.1. Un contexte d'effet Lombard

Le fait d'augmenter le niveau sonore de la voix en milieu bruyé est connu en acoustique sous le nom de réflexe Lombard (Lombard, 1911). Compte tenu des conditions d'environnement bruyé que nous avons mises en exergue, il est clair que le chœur des ABs se trouve dans une situation d'expression orale, où le réflexe Lombard va s'exercer de façon automatique. Cette élévation de la voix ne va pas être sans conséquences sur le plan acoustique, et un examen visuel rapide des spectrogrammes nous a laissé entrevoir des variations importantes qui vont affecter les trois principaux paramètres. En premier lieu, l'espacement des raies spectrales implique une élévation de la fréquence fondamentale, laquelle va avoir une incidence directe sur la fréquence des formants

---

<sup>3</sup> En maori, le groupe /wh/ (par ex. à l'initiale de *whiti te ra*) = [f or θ or hw] (Biggs, *op. cit.*, p. 470).

des voyelles. En deuxième lieu, le fait que l'intensité soit faible, si ce n'est nulle au-delà de 4.0Khz est la preuve tangible que la répartition de l'énergie spectrale est sensiblement différente de celle de la parole dite normale. Enfin, des tenues vocaliques de l'ordre de 450 ms et plus, c'est-à-dire identiques à celles que l'on trouve en voix chantée, nous renforcent dans notre opinion que nous avons effectivement affaire à une parole tout à fait inhabituelle.

## 4.2. La durée

L'examen visuel des spectres de parole montre que les sons à structure harmonique prédominent largement par rapport aux sons à structure de bruit. En réalité, la structure harmonique est quasi continue durant l'intégralité du *Haka*, et n'est rompue que de façon occasionnelle par les tenues (brèves) des occlusives sourdes ou par les silences (comparativement plus longs) des pauses qui constituent autant de 'brèches' acoustiques dans le continuum sonore. En se fondant sur ce simple constat, on peut d'ores et déjà s'attendre à ce qu'il existe des différences de durée notoires entre les voyelles et les consonnes, la longueur des premières faisant contraste avec la brièveté des secondes.

### 4.2.1. Les voyelles

Parmi les voyelles produites pendant la harangue et le *Ka mate*, on distinguera les voyelles tenues, les voyelles enchaînées ou voyelles en hiatus, les voyelles en contexte de syllabe CV, les cris assimilables sur le plan de la quantité à des voyelles de (très) longues tenues et dont la durée est souvent supérieure à une seconde.

#### 4.2.1.1. Les voyelles tenues

Celles-ci se rencontrent aussi bien en position initiale qu'en position finale. Confirmant nos premières impressions auditives, l'analyse montre que chaque vers du *Ka mate* est invariablement initié par la même voyelle [a] dont la durée est à peu près constante (entre 445 et 475ms). En revanche, la voyelle finale est variable aussi bien du point de vue de sa nature que de sa durée. En effet, les deux premiers vers du *Ka mate* se terminent par la voyelle [ɔ] dans la syllabe ouverte [kɔ:] et dont la durée est sensiblement plus longue qu'à l'initiale (> 700 ms). La voyelle finale du troisième vers est la voyelle grave [u] dont la durée est encore supérieure (1.000 ms) à celle des vers précédents. Enfin, le vers 5 se termine par une syllabe ouverte [pa] dont la voyelle est d'une durée (500 ms) sensiblement égale à celle de la voyelle initiale.

#### 4.2.1.2. Les voyelles enchaînées

En règle générale, on remarquera que l'une des deux voyelles enchaînées est TOUJOURS la voyelle [a]. Quand cette dernière constitue le premier élément de la paire, elle est au contact soit de

la voyelle antérieure [i] dans [naitiki], soit de la voyelle postérieure [u] dans [ʔikamau] (harangue) ou encore [ʔaupa] aux vers 5 et 6 du *Ka mate*. Quand l'ordre est inversé, c'est-à-dire quand [a] constitue le deuxième élément de la paire, la voyelle est plutôt associée à des voyelles antérieures, par exemple [i] dans [pakia] ou [e] dans [tea] ; dans ce dernier cas, les deux syllabes Te Ra sont réduites à une syllabe unique suite à l'amûissement de la consonne à battement(s) [r].

La durée de la deuxième voyelle [a] (enchaînée) [ʔikamau] est supérieure de 20 à 25% à celle de la même voyelle produite dans une syllabe en position initiale, cf. [karite]. Quand elle est produite à l'attaque de vers, la durée moyenne de [a] est de 455 ms, durée qui est pratiquement identique (respectivement 459 ms et 444 ms) à celle de [a] dans les deux occurrences de [karite] produites par le chef ; enfin la durée de la voyelle initiale [a] (enchaînée) dans les trois occurrences de [ʔaupa] est de 433, 457 et 425 ms, soit des durées sensiblement égales aux durées précédentes. Sur le plan de la quantité vocalique, il ressort qu'une voyelle 'type' est produite par le chef, et que ses propriétés temporelles vont être conservées, qu'elle soit voyelle initiale, voyelle finale, premier ou deuxième élément de voyelles enchaînées. Ainsi, on peut dire que la production individuelle du chef va servir de référence au chœur, dont la production collective sera en quelque sorte le calque, ou plus exactement l'écho acoustique.

#### 4.2.1.3. Les voyelles en contexte syllabique CV

Il a été mentionné plus haut que le *Haka* était perçu comme une parole 'dure' et 'scandée'. Cette scansion est particulièrement nette dans les trois premiers vers, qui sont caractérisés par la présence de six syllabes entre la voyelle initiale [a:] et la syllabe finale [kɔ:] (vers 1 et 2) ou la voyelle finale [u:] (vers 3). Dans la plupart des cas, on constate une homogénéité de durée entre les syllabes. Comme notre échantillon de données était limité, une analyse statistique n'a pu être effectuée. Cependant, si l'on tient compte de la faiblesse des écarts, (ceux-ci n'excédant pas 17%), il est plus que probable que les différences de durée n'auraient été que peu ou pas significatives.

En ce qui concerne l'influence de l'accent sur la durée vocalique, il s'avère que la voyelle perçue comme accentuée dans les syllabes 'ka' (vers 1 et 2), 'te', 'tāh' (vers 3), n'est pas nécessairement plus longue que les syllabes inaccentuées subséquentes. Au contraire, c'est la voyelle de la syllabe terminale qui est souvent la plus longue (> 200 ms), d'où un rapport V/C (de 4.50 à 5.0) qui est plus élevé pour la dernière syllabe par comparaison à la première.

**Tableau 3**  
*Durée de la voyelle, consonne, syllabe C+V et rapport de durée V/C*  
*dans les syllabes médianes des trois premiers vers du Ka mate.*

<b>Vers 1</b>	<b><i>ka</i></b>	<b><i>ma</i></b>	<b><i>tɛ</i></b>	<b><i>ka</i></b>	<b><i>ma</i></b>	<b><i>tɛ</i></b>
Voyelle	192	209	219	187	193	217
Consonne	67	44	40	66	46	38
Syllabe C+V	259	253	259	253	235	255
Rapport V/C	2.86	4.75	<b>5.47</b>	2.83	4.19	<b>5.71</b>
<b>Vers 2</b>	<b><i>ka</i></b>	<b><i>ma</i></b>	<b><i>tɛ</i></b>	<b><i>ka</i></b>	<b><i>ma</i></b>	<b><i>tɛ</i></b>
Voyelle	186	171	180	175	178	209
Consonne	64	45	50	76	51	41
Syllabe C+V	250	216	230	251	229	250
Rapport V/C	2.90	3.8	<b>4.12</b>	2.30	3.43	<b>5.09</b>
<b>Vers 3</b>	<b><i>tɛj</i></b>	<b><i>nɛ</i></b>	<b><i>tɛ</i></b>	<b><i>tã</i></b>	<b><i>ŋa</i></b>	<b><i>ta</i></b>
Voyelle	168	175	216	190	172	212
Consonne	67	43	50	68	50	48
Syllabe C+V	235	218	266	258	222	260
Rapport V/C	2.50	4.06	<b>4.32</b>	2.79	3.44	<b>4.41</b>

#### 4.2.1.4. Les cris

Comme nous l'avons remarqué précédemment, les deux cris poussés par le chœur des All Blacks sont assimilables à des voyelles tenues. Dans la harangue, le premier cri d'une durée de plus d'une seconde (1.3 sec.) a le timbre grave de la voyelle postérieure [u] et retentit comme un écho à la voyelle finale de la séquence [ʔikamau] produite par le chef. En conclusion du *Ka mate*, le deuxième cri a plutôt le timbre clair d'une voyelle antérieure et il est nettement plus bref (500 ms) que le cri précédent, cette brièveté 'relative' étant due au fait que ce cri (on pourrait également utiliser le terme de groupe de souffle terminal) est synchronisé avec le bond vertical (rapide) qui met un terme au *Haka*.

#### 4.2.2. Les consonnes

1. Contrairement aux tenues vocaliques qui sont longues (durée moyenne =187.8 ms), les tenues consonantiques sont comparativement brèves (durée moyenne =53.2ms). Avec un rapport moyen

V/C de 3.50, il apparaît que la durée des voyelles est plus de trois fois supérieure à celle des consonnes dans le seul contexte CV. De plus, si l'on fait entrer en ligne de compte la durée des voyelles initiales, des voyelle finales ou des cris, c'est-à-dire la durée de tous les segments vocaliques ou considérés comme tels, le rapport V/C s'élève à 5.05, ce qui accrédite l'impression auditive unanime des auditeurs qui est celle d'un vocalisme continu.

2. Alors que la voyelle des syllabes perçues comme accentuées, 'ka(x4), 'tej, 'tã, cf. tableau 3, n'est pas nécessairement plus longue que la voyelle des syllabes inaccentuées, c'est au contraire la durée de la consonne initiale qui se trouve augmentée. En effet, celle-ci est de 65 à 70 ms dans 'ka, 'tej, 'tã, soit une durée supérieure à celle des autres consonnes dans les syllabes subséquentes (durée moyenne = 45 ms). Dans ce contexte de tension à la fois psychique et physiologique, cet allongement de la consonne initiale accrédite la remarque de Pike (1979), selon laquelle la consonne est allongée dans de nombreuses langues du monde sous l'effet de la colère. En contrepartie, cet allongement de la tenue consonantique est compensé par un abrègement de la tenue vocalique dans la syllabe initiale, alors qu'un phénomène inverse se produit dans la syllabe terminale (tableau 3).

3. Cet effet compensatoire entre la voyelle et la consonne apparaît de façon nette (ligne 3), si l'on considère la durée globale C+V. En effet, on peut noter que les écarts de durée sont faibles entre les syllabes, et que la durée de chaque syllabe varie peu par rapport à la moyenne (245.2 ms) qui est valable pour les dix-huit syllabes concernées.

4. Les mesures acoustiques qui attestent de la brièveté des tenues consonantiques sont en contradiction avec celles effectuées par plusieurs auteurs, qui ont relevé certaines différences en fonction du recul du lieu d'articulation. C'est ce que signale Umeda à propos des consonnes de l'anglo-américain : *labials phonemes are the longest, dentals are shorter than labial and velar consonants are the shortest* (Umeda, 1977, p. 196). Pour sa part, Durand (1985) constate que la durée de la tenue décroît de /p/ à /k/ en français, opinion partagée par O'Shaugnessy (1981, p. 340), *Unvoiced labials were found longer than lingual and velar stops*. Cependant, il convient de préciser que ces résultats ont été obtenus à partir de l'analyse de mots insérés dans des phrases porteuses, donc enregistrés dans des conditions où de nombreux facteurs de variation temporelle ont été neutralisés au préalable. En d'autres termes, ces résultats concernent une parole dite de 'laboratoire', et de ce fait ne sauraient s'appliquer à une parole spontanée, et qui plus est 'criée', comme il le sera démontré par la suite. Dans le *Haka*, où la fonction primaire des consonnes est de 'hacher' le continuum sonore, la tenue se trouve abrégée et réduite au minimum, ce qui explique que l'on ne retrouve pas les micro-différences relevées ci-dessus. Cette remarque est valable pour les consonnes [k] et [t] dont la durée est pratiquement identique (67 vs 68 ms) en syllabe initiale accentuée, cf. [ka] vs [tej] et [ka] vs [tã] mais également pour la consonne labiale [p] dont la durée dans les deux occurrences de

[ʔaupɑ] (respectivement 67 et 63 ms) est pratiquement identique à celle des consonnes dento-alvéolaires et vélaires. Nos observations corroborent celles de Stathopoulos et Weismer (1983), pour lesquels ces différences de durée ne sont pas dues à un facteur unique, en l'occurrence le lieu d'articulation, mais à plusieurs facteurs comme le voisement, l'accent, le sexe ou encore la position syllabique. Dans le cas présent, il ressort que, quel que soit le lieu d'articulation de la consonne (labial, dento-alvéolaire ou vélaire) ou son mode d'articulation (voisé/non-voisé), (à ce propos, on signalera que la durée des consonnes nasales (sonores) /m, n, ŋ/ est quelquefois supérieure ou égale à celle des consonnes sourdes /p, t, k/), la durée des consonnes est la plupart du temps 'figée', et n'est sujette à des variations minimales qu'en fonction de la position initiale accentuée de la syllabe.

5. Ces différences de durée entre voyelles et consonnes nous conduisent à considérer sous un angle nouveau la question de l'homothétie entre le *Haka* et le chant. En effet, si l'on se fonde sur les résultats d'une étude comparative sur l'organisation temporelle de la syllabe en voix parlée et en voix chantée (Scotto Di Carlo et Autesserre, 1992), on constate qu'il existe, de ce point de vue, des rapports de convergence entre la voix du chœur des ABs et la voix chantée. Alors que la durée syllabique est dans une syllabe parlée, à peu près également répartie entre la voyelle et la consonne, il n'en est pas de même dans une syllabe chantée où la durée de la consonne ne représente que 12% de la durée totale de la syllabe, alors que la voyelle en représente 88%. Or, ce sont pratiquement les mêmes pourcentages (88.4% *vs* 11.6%) que l'on retrouve à propos de la durée des voyelles et des consonnes du *Ka mate* dans le cadre syllabique CV. L'allongement syllabique entraîne une modification de la durée relative des voyelles et des consonnes à l'intérieur de la syllabe, ce qui contribue à augmenter la durée vocalique au détriment de la durée consonantique. On peut donc en déduire qu'il existe de ce point de vue une organisation temporelle syllabique qui apparente la voix du chœur des All Blacks à la voix chantée.

6. En relation avec ces différences de durée entre voyelles et consonnes, un autre élément qui milite en faveur d'une assimilation à la voix chantée, est le fait que l'organisation temporelle du *Haka* est caractérisée par un dischronisme syllabique. Même si certaines syllabes médianes tendent à une égalité de durée, (*cf.* les vers 1, 2, 3 du *Ka mate*), il n'en demeure pas moins que cette égalité est due à des contraintes rythmiques qui régissent et conditionnent la production des unités phoniques. En effet, nous verrons dans le chapitre consacré aux propriétés suprasegmentales que cette rythmicité n'est pas uniforme dans le *Ka mate*, et qu'il existe des différences de durée importantes selon que les syllabes se situent en position initiale, médiane ou finale dans les vers. En tout état de cause, on peut dire que la loi de l'isochronisme syllabique, selon laquelle l'équilibre est maintenu dans la parole dite 'normale' par les rapports de durée entre les voyelles et les

consonnes, ne s'applique pas à un type de parole comme le *Haka*, qui est caractérisé foncièrement par un déséquilibre temporel à l'intérieur de la syllabe.

En résumé, nous dirons que, s'il existe une analogie entre la voix du *Haka* et la voix chantée du point de vue des rapports de durée entre voyelle et consonne, c'est que, dans un cas comme dans l'autre, *l'équilibre temporel syllabique se trouve détruit* (Scotto et Autesserre, p. 10, 1994).

#### 4.3. L'intensité

Dans une section précédente (1.3.2), nous avons insisté sur l'état de tension qui régnait parmi des joueurs avant l'exécution du *Haka*. Cet état de tension psychologique va se manifester sur le plan physiologique par une tension du mécanisme laryngé et par une surpression sous-glottique. Préalablement à l'acte de phonation, les cordes vocales sont tendues et fermées, et l'expulsion brutale de l'air pulmonaire va entraîner la mise en vibration instantanée des cordes vocales. En se fondant sur l'examen de radiographies du larynx, mais aussi sur l'observation d'images laryngoscopiques des ventricules et des cordes vocales dans diverses productions vocales, Fónagy avait observé que l'excitation ou la colère se manifestaient au niveau glottique par une contraction spasmodique des muscles laryngés, *les cordes vocales sont fortement accolées, contractées, les bandes ventriculaires s'imbriquent et les ventricules de Morgani sont encore plus rétrécis* (Fónagy, *op. cit.*, p. 45-46). En fait, la glotte se transforme en un canal étroit, ce qui va favoriser la formation d'une occlusion glottique, laquelle se rencontre régulièrement à l'initiale vocalique du mot ou de la phrase dans les émissions sonores agressives. Comme le *Haka* correspond précisément à ce type de message, on constate que les voyelles [a] en position initiale de vers ou en tant que premier élément de voyelles enchaînées, sont produites avec un coup de glotte pré-vocalique [ʔ] dont le corrélat acoustique est une montée abrupte de l'intensité. La réalité de cet indice est mise en évidence sur la figure 1, où l'on observe que la voyelle initiale [a] est caractérisée par une attaque brutale de l'intensité. Compte tenu du fait que le système phonatoire se trouve à la limite du fonctionnement, la courbe d'intensité atteint rapidement son amplitude maximale. Elle se stabilise ensuite, et prend la forme d'un plateau avec une chute (minime) de 3 dB sur la syllabe terminale. En réalité, le niveau d'intensité maximale (de l'ordre de 50 à 55dB) tel qu'il est affiché sur le document (figure 1, C), n'est pas représentatif du niveau d'énergie rayonnée aux lèvres des locuteurs. En effet, ce niveau est anormalement bas, puisqu'il est inférieur de plus de 10 dB à celui d'une voix parlée normale (65 dB), et de près de 30 dB par rapport à celui d'une voix criée (niveau moyen 93-95 dB), la différence généralement admise entre ces deux types de voix étant de 28 dB. À notre opinion, cette disparité entre l'intensité affichée et l'intensité réelle s'explique de la façon suivante :



En premier lieu, il convient de garder à l'esprit que le son est enregistré en champ libre et en milieu bruité, c'est-à-dire dans des conditions qui n'ont rien à voir avec les conditions optimales qui sont celles d'un studio d'enregistrement ou d'une chambre sourde.

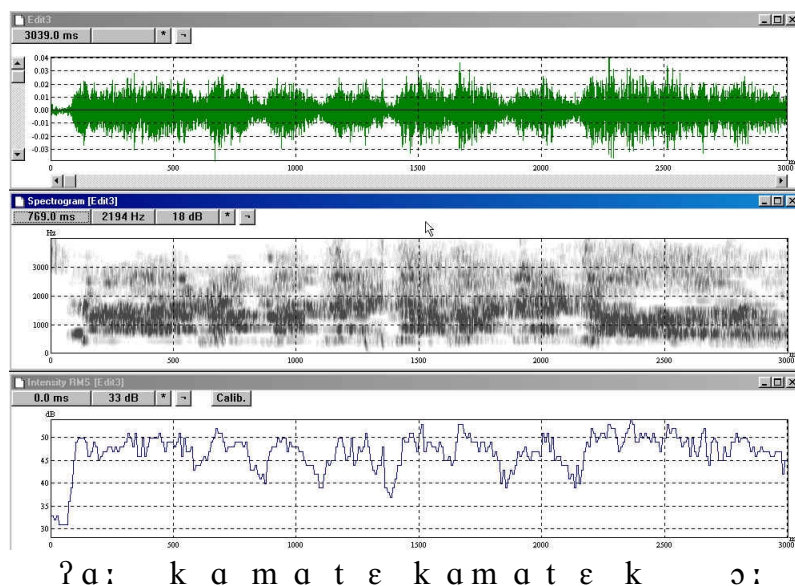
En deuxième lieu, la production globale émane non pas d'une voix, mais de vingt-deux voix enregistrées par l'intermédiaire de microphones qui ont chacun leurs caractéristiques techniques en termes de largeur de bande passante, et qui plus est, sont positionnés à distance variable des joueurs/locuteurs. C'est ainsi que le microphone placé devant la bouche du chef est maintenu au bout d'une perche, elle-même tenue à bout de bras par un preneur de son ; en réalité, le microphone n'est jamais fixe *stricto sensu* et est donc sujet à des déplacements, si minimes soient-ils. Les microphones placés sur des supports devant le chœur sont fixes, mais se situent à une plus grande distance des joueurs (2 à 3 mètres), lesquels ne sont pas toujours statiques puisqu'il leur arrive de se déplacer vers l'adversaire pendant l'exécution du *Haka*. Enfin, il est possible que le son soit enregistré par l'intermédiaire des microphones intégrés dans les caméscopes des cameramen qui se déplacent eux-mêmes pour filmer les joueurs sous un angle de vue différent. De la distance variable entre l'émetteur du son (le joueur/locuteur) et le récepteur (le microphone du preneur de son ou du cameraman), va résulter une variation non seulement de l'intensité du son, mais aussi du rapport signal/bruit. Enfin, il convient de tenir compte d'un autre facteur de variation, lié à l'intervention du régisseur de son. Comme les joueurs ont tendance à 'forcer' leurs voix, ce dernier va être amené à baisser le niveau sonore du signal d'entrée afin d'atténuer les montées d'amplitude du son et éviter ainsi une saturation du signal. En résumé, deux facteurs expliquent les raisons pour lesquelles le niveau d'intensité affichée est bas par rapport à celui de l'intensité réelle ; d'une part le facteur 'spatial', c'est-à-dire la distance (variable) entre le locuteur et le microphone, d'autre part le facteur 'humain' dû à l'action du régisseur qui va 'moduler' le niveau du signal d'entrée afin de placer le téléspectateur dans les meilleures conditions d'écoute possibles.

En règle générale, la courbe d'intensité se caractérise par sa stabilité et sa continuité, mais cette continuité 'apparente' n'en est pas moins rompue en plusieurs occasions. Ces ruptures sont dues aux consonnes occlusives /k/ et /t/ et (dans une moindre mesure) à la consonne nasale /m/, ces deux catégories de consonnes étant caractérisées par une fermeture complète du conduit vocal, par exemple dans la séquence *Ka mate*. Toutefois, ces ruptures de 5 dB à 10 dB (selon la nature des consonnes) ne constituent pas un indice pertinent pour la perception des segments consonantiques dans un environnement bruité quel qu'il soit. Telle semble être l'opinion de Rostolland (1982) pour qui la voix criée<sup>4</sup> entendue à distance ou noyée dans le bruit, ne semble contenir que des voyelles, les indices

---

<sup>4</sup> Ceci s'applique également à la perception de la voix chantée en présence d'un orchestre de 80 musiciens ou plus, remarque Nicole Scotto Di Carlo.

acoustiques des consonnes n'étant pas suffisamment robustes pour en permettre la perception, et à plus forte raison la discrimination.



**Figure 1**  
*Oscillogramme (A), spectrogramme (B) et courbe d'intensité globale (C)*  
*du premier vers du Haka [ʔa:kamatekamatekɔ:]*

#### 4.4. La fréquence

##### 4.4.1. La fréquence fondamentale

Afin de faire en sorte que sa voix porte le plus loin possible, le locuteur va augmenter la pression sous-glottique. Comme il tend et comprime simultanément les cordes vocales, l'amplitude de la vibration glottique augmente, mais également sa fréquence. Cette élévation de la fréquence fondamentale (F0) est mise en évidence sur la figure 2 où la montée atteint une fréquence supérieure à 400 Hz. Cette montée abrupte est suivie d'un plateau mélodique d'une durée de 2.5 sec. environ, avant que la courbe ne s'incurve légèrement dans sa partie terminale sur la voyelle [ɔ:]. L'existence de ce plateau intonatif s'explique du fait qu'une fréquence fondamentale maximale est atteinte dès la première impulsion glottique. En conséquence de quoi, et à partir du moment où la limite supérieure de la tessiture du ou des locuteurs est atteinte, la courbe de F0 ne peut plus

s'élever, mais seulement se stabiliser ou s'infléchir vers le bas. Comme la courbe d'intensité, la courbe mélodique est caractérisée par la présence de ruptures ; toutefois, celles-ci sont de plus forte amplitude, ce qui s'explique de la façon suivante.

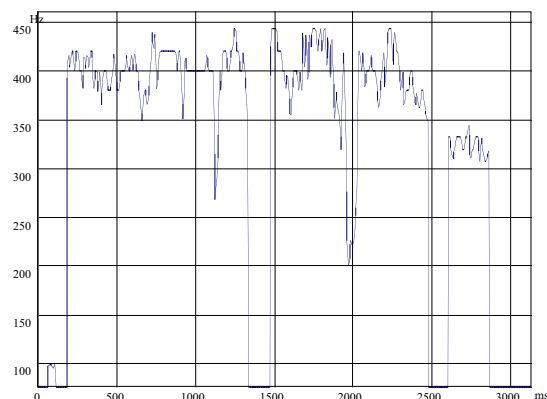
1. À la différence de la harangue qui est relativement discontinue en raison de plusieurs pauses qui sont observées à la fois par le chef et par ses équipiers qui répondent à ses exhortations, le *Ka mate* est produit de façon plus continue. Néanmoins, si on observe attentivement la figure 2, on constate que la courbe mélodique s'infléchit de façon abrupte vers le bas en plusieurs occasions. Si une rupture de 80 à 100 Hz s'explique sur le plan acoustique, cf. la syllabe *Ka* à  $T = 665$  ms.) par les variations micro-mélodiques qui accompagnent habituellement le passage d'une consonne occlusive à la voyelle contiguë (Di Cristo, 1985), par contre des ruptures de 200 Hz à  $T = 1130$  ms (syllabe *te*) ainsi qu'à  $T = 2000$  ms (2<sup>e</sup> syllabe *te*), ne sauraient s'expliquer en relation avec de tels phénomènes micro-prosodiques. Plus encore, les décrochages à  $T = 1500$  ms,  $T = 2500$  ms et  $T = 3000$  ms, où la  $F_0$  chute de 300 Hz jusqu'à la fréquence zéro, ne trouvent pas d'explication logique en fonction d'une influence consonantique quelconque. C'est le cas à  $T = 2500$  ms. où la chute de  $F_0$  est suivie d'une remontée vers un plateau intonatif à 325 Hz. Après s'être stabilisée, la fréquence chute à nouveau en dépit du fait que le segment phonique en question soit voisé, en l'occurrence la voyelle [ɔ:], comme en témoigne la présence de vibrations laryngées de forte amplitude et une structure de formants continue (figure 1). De toute évidence, l'origine de ces décrochages de  $F_0$  doit être recherchée dans les conditions d'enregistrement du signal acoustique et du mauvais rapport signal/bruit qui en résulte. Si l'utilisation du logiciel *Sound Forge* a permis de réduire de 10 dB le niveau du bruit ambiant, il est certain que cette réduction a été insuffisante pour neutraliser les artéfacts et/ou les bruits parasites qui accompagnent l'émission du message verbal. En tout état de cause, si les logiciels *Phonedit* et *Praat* ont permis de détecter et de visualiser la  $F_0$ , il n'en demeure pas moins que l'obtention d'une courbe mélodique exempte de tout décrochage relevait de l'impossible compte tenu des raisons invoquées ci-dessus.

2. La hauteur moyenne de  $F_0$  (405 Hz), calculée pour l'ensemble des voyelles du *Haka*, est élevée par rapport à la hauteur moyenne de voix masculines (150 Hz) ou encore de voix féminines (295 Hz). Cette fréquence élevée qui se situe à l'extrême limite de la tessiture des locuteurs, s'explique comme nous l'avons mentionné par une surtension du mécanisme laryngé. Elle est en fait LE critère acoustique primaire qui permet de différencier formellement la voix criée de la voix parlée. En effet, si l'on applique à la  $F_0$  d'une voix masculine normale (150 Hz) le facteur multiplicatif de 2.5 ou 3 préconisé par Lafon (1961), et si l'on tient compte du fait que la fréquence de  $F_0$  d'une voix criée augmente en moyenne de 2 octaves par rapport à la voix parlée, on obtient des

fréquences qui s'échelonnent de 375 à 450 Hz, c'est-à-dire la gamme fréquentielle où se situent précisément la plupart des fréquences de F0 mesurées sur les voyelles du *Haka*.

3. En ce qui concerne les variations tant ascendantes que descendantes de F0, leur étendue est des plus réduites. En règle générale, et ce fait s'est avéré vérifié aussi bien dans la harangue que dans le *Ka mate*, la fréquence d'attaque de la voyelle initiale [a] se situe invariablement au-dessus de 400 Hz, hauteur qui est stable durant la plus grande partie du vers. La courbe mélodique du vers 1 (figure 2) est à cet égard représentative de celle des autres vers du *Ka mate*, en ce sens que leurs contours mélodiques respectifs s'avèrent être tous à peu près identiques. En effet, la ligne mélodique est uniformément plate pendant toute la durée du vers, soit pendant plus de 2500 ms (exception faite des décrochages dont nous avons expliqué la présence) et ne s'infléchit légèrement que dans sa partie terminale. Toutefois, cette chute reste modérée et n'excède pas 80 Hz, soit moins d'un quart d'octave. Le maintien d'une fréquence fondamentale élevée est favorisé à la finale par la production soit d'une voyelle de longue durée, en l'occurrence [ɔ:] aux vers 1 et 2), [u:] au vers 3, [ɑ:] au vers 5, soit de voyelles enchaînées, [eɑ] au vers 4, ou encore par un long cri dont le timbre aigu l'assimile à la voyelle antérieure [i:] au vers 6. Nous aurons l'occasion de revenir sur les points de similitude entre la musique et le *Haka*, mais dès à présent, on remarquera que les paramètres durée et fréquence évoluent de concert, c'est-à-dire qu'à une stabilité temporelle correspond une stabilité fréquentielle. De ce point de vue, la production du locuteur est ici similaire à celle du chanteur, dont la note finale serait tenue dans un registre élevé à la fin d'un refrain ou d'une chanson. C'est le cas de la voyelle finale [u:] dans la séquence [ʔikamau] dont les modulations de F0 n'excèdent pas 30 Hz sur toute la durée de la tenue vocalique (800 ms), ou encore de la voyelle finale [i:] dans [ʔakakino # ʔaki:] dont les variations extrêmes sont du même ordre pour une tenue d'une durée supérieure à la seconde. Le même constat s'impose à propos de la voyelle finale [u:] du vers 3 (>1.5 sec.) et dont la fréquence d'arrivée (400 Hz) est pratiquement égale à la fréquence d'attaque (406 Hz). Enfin, l'écart de 52 Hz entre la première et la dernière impulsion glottique du cri final (412-380 Hz) atteste qu'une F0 élevée et stable a été maintenue durant la majeure partie de la tenue vocalique.

4. Si l'amplitude des variations mélodiques descendantes est faible, il en est de même de leurs homologues ascendantes. En effet, si l'on considère que le phénomène de saturation fréquentielle est occasionné par une tension maximale et une masse vibrante minimale des cordes vocales (Husson, 1960), la F0 ne peut avoir qu'une marge d'élévation limitée. La seule montée sensible concerne la voyelle finale [i:] (450 ms) de la séquence [majwajtri:] produite par le chef, dont la F0 culmine à 482 Hz, pic qui constitue en fait le sommet mélodique du *Haka*. Compte tenu de son timbre clair, la voyelle est perçue comme un son strident émis par une voix aiguë (... a *high-sbrilled voice*) laquelle tran-



**Figure 2**  
*Courbe mélodique du premier vers [ʔa'kamate'kamate'kɔ:]*

che avec le timbre grave des voix mâles du chœur, voix qui n'est pas sans rappeler celle de John Timu (un autre chef de *Haka*, ... *John's high-pitched scream of Haka exertion*, K. Quinn, *op. cit.*).

5. Enfin, on ajoutera que l'impression qui se dégage du chœur est celle d'une grande uniformité des voix. Ce phénomène d'osmose vocale n'est pas surprenant en soi, si l'on considère que le fait de crier fait perdre à la voix une grande partie de sa spécificité, et qu'une élévation de F0 jusqu'à 400 Hz atténue, si ce n'est annihile totalement les caractéristiques idiosyncrasiques du locuteur. On remarquera à ce sujet que cette uniformité de timbre se retrouve dans la voix chantée du fait que le tractus vocal est tétanisé, crispé, contracté, figé<sup>5</sup>. Dans le cadre d'un travail de recherche sur les propriétés vocales des locuteurs, il serait intéressant de voir dans quelle mesure les différences intrinsèques de F0 s'annulent en voix criée ou si elles sont au contraire préservées comme c'est le cas en voix parlée, où la hauteur de F0 est fonction du degré d'ouverture de la voyelle.

#### 4.5. La relation entre propriétés articulatoires, aérodynamiques et acoustiques

Grâce à la technique magnéto-métrique (Schulman, 1988), il a été démontré sur le plan physiologique que la voix criée différait de la voix parlée par une augmentation de l'amplitude des mouvements de la mâchoire. Sur ce plan, les effets 'naturels' en parole criée sont proches des effets 'artificiels' provoqués par un blocage de la mâchoire (bite-block speech), dans la mesure où

---

<sup>5</sup> Les termes sont de N. Scotto Di Carlo.

ils entraînent tous deux une augmentation du degré d'ouverture de l'angle mandibulaire. Cette augmentation de l'angle mandibulaire entraîne conjointement une augmentation de l'aperture dans le tractus vocal, et inversement une diminution de la taille de la constriction. Ces changements de configurations articulatoires s'expliquent en fonction de phénomènes aérodynamiques.

Dans une situation d'environnement bruyé, le locuteur va naturellement augmenter la pression sous-glottique, laquelle va augmenter à son tour la pression supra-glottique, ainsi que le flux d'air dans la cavité orale. Si le locuteur essaye de maintenir en voix criée une articulation dite 'normale' pour les voyelles fermées [i] ou [u], c'est-à-dire une constriction étroite du conduit vocal, l'augmentation du flux d'air et la baisse de pression au niveau de la constriction vont avoir pour effet de provoquer l'émission d'un bruit turbulent comparable à celui des consonnes fricatives, mais qui est '*sur le plan phonologique inacceptable pour des segments vocaliques*' (Schulman, *op. cit.*, p. 310).

La relation entre flux d'air, taille de la constriction et chute de la pression à la constriction a été exprimée par Stevens (1971) en fonction de l'équation  $\Delta P = k (\rho U^2 / 2A^2)$ , où  $\Delta P$  est la chute de la pression à travers la constriction,  $U$  la vitesse volumique de l'air, et  $A$  la surface transversale de la constriction.

Selon cette équation, si l'on tient compte de l'accroissement du flux d'air à la constriction, une augmentation de la surface de ladite constriction est rendue nécessaire pour empêcher l'émission d'un bruit turbulent. En nous fondant sur les fréquences de formants, nous verrons plus loin que cette condition est remplie pour les voyelles fermées [i] et [u] caractérisées toutes deux par une diminution de la constriction occasionnée par un abaissement de la masse linguale ; ce mouvement, qui a pour effet d'augmenter la surface d'aire au lieu de constriction, s'accompagne respectivement d'un recul de la langue pour [i] et d'une avancée de cette même masse linguale pour [u]. Afin de prévenir une confusion articulatoire (et perceptive) avec les voyelles mi-fermées [e-o] ou mi-ouvertes [ɛ-ɔ], il va se produire un effet de domino, par lequel les voyelles criées vont être articulées avec une ouverture accrue de la mâchoire, ce qui occasionnera à son tour une position abaissée de la masse linguale.

## 4.6. Les propriétés spectrales des voyelles

### 4.6.1. L'enveloppe spectrale

Si l'on s'appuie sur les (rares) données acoustiques recueillies à propos de la voix criée, il semble que le cri déforme la structure acoustique des sons dans le triple espace temps, énergie, fréquence. En ce qui concerne ce dernier paramètre, il est inéluctable qu'une élévation de la fréquence fondamentale par un facteur de plus de 2 octaves ne peut avoir que des conséquences considérables sur les propriétés spectrales des sons.

1. L'une des conséquences les plus évidentes, et au sens littéral du terme les plus 'visibles' de cette élévation de F0, est l'espacement des raies spectrales sur un spectrogramme de parole. Cet écartement des harmoniques, qui n'est pas sans rappeler celui des voix féminines ou enfantines, ne facilite pas la mesure de la fréquence des formants. C'est pourquoi une double analyse par Transformée de Fourier rapide (FFT) et par prédiction linéaire (LPC) a été effectuée pour minimiser les risques d'erreur dans la détection et la mesure des pics spectraux.

2. Une autre conséquence de l'élévation de F0 a trait aux différences entre les composantes à basse et à haute fréquence, différences qui se reflètent dans la forme de l'enveloppe spectrale. Alors que les voyelles 'parlées' non soumises à l'effet Lombard sont caractérisées par une chute rapide de l'énergie après le pic qui englobe les harmoniques de F0 et de F1, les voyelles 'criées' du *Haka* se différencient des précédentes par une chute moins abrupte de l'énergie, et un spectre plat dû à l'étalement de l'énergie entre 0.4 KHz et 4.0 KHz. Cette différence en termes de forme de l'enveloppe spectrale, implique un transfert de l'énergie vers les fréquences moyennes au détriment des fréquences basses.

3. Ce déplacement de l'énergie, qui sous-entend une modification de la pente d'inclinaison spectrale (*spectral tilt*), est généralement considéré comme un indice acoustique primaire des voyelles produites en milieu bruyant (Van Summers *et al.*, 1988). Quoique cet indice n'ait pas fait l'objet d'une analyse particulière dans le cadre de cette étude, nous avons pu recueillir indirectement des informations sur la pente d'inclinaison spectrale par la mesure du niveau d'énergie dans les fréquences basses du spectre. C'est ainsi qu'une chute de 6 dB par rapport à des voyelles parlées 'normales', a été constatée dans la gamme 0-400 Hz, écart qui correspond *grasso modo* à la chute rapportée par Junqua (1996). Cette chute d'énergie a pour conséquence une pente moins accentuée de la ligne d'inclinaison spectrale, laquelle entraîne conjointement un aplatissement de l'enveloppe spectrale dans la gamme concernée. Ainsi, on peut dire que le déplacement de l'énergie vers cette zone est LA caractéristique acoustique des voyelles criées.

En résumé, les voyelles du *Haka* se caractérisent par un niveau d'énergie comparativement faible entre 0 et 400 Hz, une forte concentration entre 400 Hz et 4000 Hz<sup>6</sup>, et une absence quasi totale d'énergie à

---

<sup>6</sup> C'est précisément aux alentours de cette fréquence que se situe le *singing formant* des chanteurs, lequel est responsable de la portée de la voix et permet à ces derniers de dominer la masse orchestrale. Ce renforcement entre 3000 et 5000 Hz se retrouve aussi dans le spectre du signal sonore de nombreux instruments de musique à longue portée (clairons, trompettes) lesquels pouvaient couvrir le fracas des combats. Le chant du coq qui présente cette même caractéristique porte aussi très loin (N. Scotto di Carlo).

*Note personnelle de l'auteur* : il se peut que le coq (gaulois) importé de France à l'occasion, soit introduit dans les stades néo-zélandais par des supporters (gaulois évidemment) pour contrecarrer l'impact sonore du *Haka*, cf. Le chanteur et l'orchestre : David contre Goliath ?, *La lettre du musicien*, 2001, 249, p. 10.

partir de cette dernière fréquence jusqu'à 5000 Hz. Toutefois, il convient de reconnaître que nous ne disposons d'aucune information concernant la largeur de bande passante des différents appareils enregistreurs (microphones fixes ou mobiles, magnétoscopes etc.) utilisés par les techniciens du son pour l'enregistrement du *Haka*. Toutefois, on signalera que Stanton *et al.* (1989) ont constaté une distribution de l'énergie spectrale similaire à propos de voyelles produites dans des conditions d'effet Lombard, où le sujet était exposé à un bruit rose de 90 dB transmis par l'intermédiaire d'écouteurs.

#### 4.6.2. La relation acoustico-articulaire

Si l'on essaye d'établir une corrélation entre les niveaux acoustique et articulaire, c'est-à-dire entre la position des formants sur les axes F1/F2 et la position de la langue sur les deux axes haut/bas et antérieur/postérieur, la figure 3 permet de constater que deux tendances se dégagent nettement :

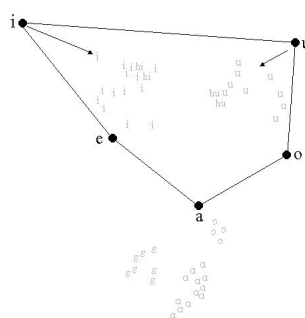


Figure 3

*Trapèze acoustique illustrant les différences de fréquence des formants 1 et 2 entre les voyelles 'criées' (en noir) du chœur du Haka et les voyelles 'parlées' (en gris) d'un locuteur maori ; les symboles hi et hu concernent les tenues vocaliques assimilées à des cris de timbre différent*

- une première tendance à l'abaissement qui se traduit par un décalage vers le bas des voyelles criées par rapport aux voyelles parlées, tendance qui laisse entrevoir une position comparativement basse du corps de la langue, laquelle est proportionnelle à l'ouverture de l'angle maxillaire. Cet abaissement de la langue qui a pour corrélat acoustique un accroissement de F1, atteste que les voyelles criées du *Haka* sont TOUTES caractérisées par une augmentation de l'aperture buccale.
- une deuxième tendance à la centralisation des voyelles criées par rapport à leurs homologues parlées, et qui n'est pas sans rappeler celle propre à la réduction vocalique



(Delattre, 1965). Toutefois, alors que ce dernier phénomène se manifeste par un mouvement convergent de TOUTES les voyelles vers une voyelle centrale [ə], ce mouvement centralisateur ne concerne pas les voyelles criées au même titre. En effet, si l'on se place sur un axe vertical, on remarque que seul l'agglomérat des points de la voyelle criée [a] se situe à l'aplomb de la voyelle parlée correspondante. D'un point de vue articulatoire, cela signifie qu'en voix criée comme en voix parlée, la configuration du conduit vocal est demeurée à peu près identique pour cette voyelle, et que le lieu d'articulation est resté à peu près identique, même si l'élévation de la fréquence de F1 peut être interprétée comme l'indice d'une constriction pharyngale plus marquée<sup>7</sup>. Par contre, les autres voyelles criées [i-ɛ-ɔ-u] ne se situent pas à la verticale des voyelles parlées correspondantes, mais se trouvent décalées vers le centre. En se fondant sur la position des fréquences de formants, on peut en déduire que sur un plan articulatoire, les voyelles criées antérieures [i-ɛ] et postérieures [u-ɔ] ont été respectivement produites avec un recul<sup>8</sup> et une avancée du corps de la langue, déplacement qui les éloignent de la périphérie et les distinguent de ce point de vue de leurs homologues parlées. De plus, si les voyelles fermées [i-u] sont caractérisées par un mouvement descendant vers le centre du trapèze, tel n'est pas le cas des voyelles ouvertes [ɛ-ɔ-ɔ]. En effet, celles-ci ne témoignent d'aucun mouvement ascendant qui serait indicateur d'une tendance à la fermeture du conduit vocal ; au contraire, elles se situent résolument à l'extérieur du trapèze, et non pas à l'intérieur de celui-ci, ce qui atteste à l'évidence d'une aperture accrue due à des conditions d'élocution et à un type de production orale tout à fait hors normes.

En conclusion de cette tentative de corrélation acoustique à visée articulatoire, il apparaît que les voyelles 'criées' se distinguent de leurs homologues 'parlées' par une plus grande aperture, et un déplacement de la langue vers une position plus centralisée. Toutefois, l'amplitude de ce déplacement est fonction de la nature de la voyelle, comme en témoigne le comportement différent des voyelles fermées [i-u] par rapport à la voyelle ouverte [a].

---

<sup>7</sup> C'est ce que les données acoustiques recueillies par Mc Lagan *et al.* (2004b) tendent à démontrer concernant l'opposition [a] vs [a:] en maori, où l'augmentation de la durée vocalique entraîne conjointement une augmentation de l'aperture et une pharyngalisation de la voyelle.

<sup>8</sup> Dans le chant, un recul du lieu d'articulation a été observé pour les voyelles antérieures, ce qui est un signe de sous-articulation.

#### 4.6.3. Les formants

Le trapèze acoustique de la figure 3 illustre la position des formants F1 et F2 des voyelles ‘criées’ du *Haka* par rapport aux voyelles ‘parlées’ d’un locuteur maori d’un âge indéterminé, mais probablement assez avancé. L’enregistrement de ce locuteur est extrait d’un document sonore, *Owing the Haka*, de Tarek Bazley (octobre 2000), et le *Ka mate* a été produit en dehors de son contexte rugbystique habituel, c’est-à-dire que le signal de parole est affranchi des interférences acoustiques qui accompagnent habituellement le *Haka* exécuté *in situ*. Sur le plan perceptif, la voix du locuteur peut être considérée comme une voix ‘parlée’, quoiqu’une légère montée mélodique soit perçue à la fin des vers 5 et 6. Malgré cette légère variation de F0, la voix de ce locuteur diffère foncièrement de celle du chef, et à plus forte raison de la voix qui émane du chœur des All Blacks. À ce titre, elle offre une solide base de comparaison entre une voix ‘parlée’ et une voix ‘criée’ ou *a fortiori* chantée.

##### Le premier formant

Comme nous avons pu l’observer, les voyelles criées se situent à une position plus basse que les voyelles parlées sur l’axe de F1, ce qui signifie que la fréquence moyenne des premières, par exemple [i] (436 Hz) et [u] (427 Hz), est supérieure à celle des secondes dont la fréquence moyenne avoisine 300 Hz dans les deux cas. Toutefois, la dispersion des voyelles témoigne d’une variabilité importante en fonction des contraintes contextuelles auxquelles elles sont soumises. En effet, la cible articuloire canonique d’une voyelle aura plus de chances d’être atteinte en contexte de voyelle isolée qu’en contexte consonantique (Stevens & House, 1963). Une variation de 100 Hz du F1 de la voyelle finale dans la séquence [ʔakinonehoki:] implique que ladite voyelle a été produite en tant que variante plus ouverte que la même voyelle [i] en contexte de syllabe [ki] ou en contexte bi-syllabique [tiki] (vers 4), l’influence des consonnes occlusives étant dans ces deux cas une influence de ‘fermeture’. Cette remarque s’applique pareillement à la voyelle [u] dont les fréquences extrêmes de F1 sont 365 et 462 Hz, ce qui laisse supposer une position haute de la langue dans la séquence [kupane] (vers 6), et *a contrario* une position plus basse en position terminale [u: # u:] (vers 3). De même, si on observe la position des nuages de voyelles, des différences apparaissent entre les voyelles [i-u] et les voyelles [ɛ-ɑ-ɔ]. En effet, alors que les premières se situent à l’intérieur du trapèze acoustique, les secondes se situent à l’extérieur de celui-ci, et plus exactement à son extrémité inférieure. En relation avec les variations de F1 rapportées ci-dessus, on constate que si les voyelles [i-u] sont relativement dispersées, les voyelles [ɛ-ɑ-ɔ] forment en revanche un agglomérat beaucoup plus compact, ce qui témoigne d’une variabilité moindre des secondes par rapport aux premières. En ce qui concerne l’opposition voyelles parlées/voyelles criées, les différences sont considérables entre les voyelles parlées [e-o] et les voyelles criées [ɛ-ɔ] dont l’aperture est proche de celle de la voyelle parlée [ɑ]. En

effet, les mesures de F1 de [ɛ] et [ɔ] révèlent des différences > à 200 Hz par rapport aux voyelles parlées [e] et [o], ce qui confirme la tendance à l'ouverture des voyelles intermédiaires du maori contemporain (Harawira, *op. cit.*), tendance qui se trouve naturellement accrue en voix criée. L'ouverture de ces voyelles fait que les fréquences moyennes de [ɛ] (763 Hz) et de [ɔ] (736 Hz) sont élevées et proches de celles de [ɑ] (806 Hz). Le fait que ces fréquences soient sinon identiques du moins voisines, confirme les données acoustiques de Rostolland (*op. cit.*) qui indique pour sa part qu'une fréquence de F1 de 700 Hz est globalement applicable à ces trois voyelles en français. Enfin, la figure 3 montre que les cris poussés pendant la harangue et à la fin du *Haka*, n'ont pas la même couleur vocalique. Le premier cri, qui est en quelque sorte la continuation de la dernière voyelle produite par le chef dans la séquence [ʔikamau] peut être assimilé à une voyelle de timbre grave. En revanche, tel n'est pas le cas du deuxième cri produit par le groupe à la fin du *Haka*, et dont le timbre est nettement plus clair comme en atteste une position plus antérieure sur le trapèze acoustique.

### Le deuxième formant

Compte tenu des écarts minimes de F1, on peut s'attendre à ce que ce soit la fréquence du deuxième formant (F2) qui joue le rôle distinctif primaire dans la distinction vocalique. Le tableau 3, où sont recensées les fréquences moyennes des trois premiers formants des voyelles criées, illustre l'importance du F2. En effet, alors que les écarts de fréquence de F1 entre les voyelles fermées [i-u] et les trois voyelles mi-ouvertes ou ouvertes [ɛ-ɑ-ɔ] ne permettent pas de distinguer ces voyelles entre elles (ceci étant également vrai pour le F3), c'est la fréquence régulièrement décroissante de F2 de [i] à [u] qui constitue effectivement l'indice de distinction primaire entre ces voyelles.

**Tableau 3**  
*Fréquences moyennes des formants F1, F2, F3 des voyelles criées du Haka  
produites par le chœur des All Blacks (22 voix mâles)*

	/i/	/ɛ/	/ɑ/	/ɔ/	/u/
F1	436	763	806	746	427
F2	1796	1598	1348	1245	1107
F3	2374	2238	2223	2155	2108

Toutefois, le seul critère relatif à la fréquence de F2 est insuffisant pour rendre compte des différences de déviation acoustique des voyelles criées par rapport aux voyelles parlées. Afin de bénéficier d'informations complémentaires, les différences ont été calculées en termes de distance

euclidienne entre les voyelles par application de la formule indiquée ci-dessous, où c = voix criée et p = voix parlée.

$$\sqrt{(F1c - F1p)^2 + (F2c - F2p)^2}$$

Les calculs montrent que la distance qui sépare les voyelles criées [i] (558.18), [u] (282.167), [ɛ] (324.54) et [ɔ] (377.68) de la voyelle parlée correspondante est nettement supérieure à celle qui sépare la voyelle criée [a] de son homologue parlée (110.32). Ainsi se trouve confirmée en termes de données objectives, l'impression visuelle qui se dégage de la figure 3, où les nombreuses réalisations de la voyelle criée [a] se situent à proximité de sa correspondante parlée, ce qui n'est pas le cas des autres voyelles, cette proximité s'expliquant sur le plan articulatoire par une contrainte d'aperture maximale au niveau buccal.

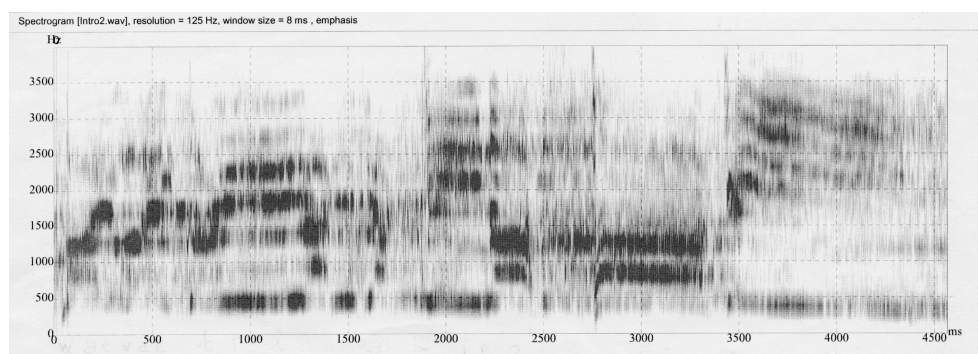
### Le troisième formant et les formants supérieurs

En l'état de la recherche acoustique, peu de données quantitatives sont actuellement disponibles concernant les formants F3, F4 ou F5 des voyelles criées. Selon Rostolland (*op. cit.*), le troisième formant des voyelles du français est relativement invariant, à l'exception du F3 des voyelles [i] et [u] dont la fréquence est respectivement supérieure (3000 Hz) et inférieure (1950 Hz) à la fréquence moyenne (2250 Hz) qui s'applique au F3 des autres voyelles. Dans leur ensemble, nos données vont dans le sens des siennes, quoique la fréquence de [i] (2.6 KHz) du chef du *Haka* soit inférieure à la valeur moyenne de F3 (3000 Hz) indiquée par Rostolland (*op. cit.*, p. 120), alors que la fréquence de [u] (2300 Hz) est au contraire plus élevée en maori par rapport à la fréquence moyenne (1950 Hz) du F3 des locuteurs français. Cette divergence de résultats ne peut être considérée que comme normale, en raison des différences de timbre sous-jacentes aux systèmes vocaliques des deux langues, mais aussi en raison d'un échantillon de parole des plus réduits.

Toutefois, plus que les variations de fréquences des formants, variations qui trouvent leur explication dans les différences de langues ou de locuteurs, c'est la concentration des formants dans une zone fréquentielle déterminée qui doit retenir l'attention. En premier lieu, on constate que la majeure partie de l'énergie acoustique se trouve concentrée sur le spectrogramme dans la gamme 400-4000 Hz, et qu'au-delà de cette dernière fréquence, l'énergie est pratiquement nulle. En deuxième lieu, la concentration des formants est apparente aux temps T = 2000ms et T = 4000ms (dans chaque cas, la voyelle concernée est la voyelle [i]), ce qui explique la difficulté que nous avons eue de mesurer la fréquence des formants supérieurs, étant donné que trois ou quatre formants se trouvaient dans une gamme inférieure ou égale à 1000 Hz (figure 4). Cette concentration explique pourquoi nous avons

eu recours à une procédure de mesure manuelle pour mesurer la fréquence de F3, ceci afin de pallier les insuffisances d'analyse automatique par FFT ou par LPC. En effet, celle-ci ne donnait lieu le plus souvent qu'à la détection d'un pic unique dont la fréquence était égale à la moyenne pondérée de pics de fréquence voisine. À ce propos, on signalera qu'en voix chantée, Sundberg (1974) a constaté le même regroupement de F3, F4, F5 en un pic unique 'intense' appelé 'formant de chant' (*singing formant*), phénomène acoustique qu'il attribue conjointement à un élargissement du pharynx et à un abaissement du larynx sur le plan physiologique.

Enfin, on précisera qu'en psycho-acoustique, il a été démontré que ce conglomerat de formants s'expliquait par l'existence d'une zone de fréquence centrée autour de 4000 Hz, zone où la voix criée est 70 fois plus intense que la voix parlée, bien que globalement parlant, la première ne soit que 25 fois plus intense que la seconde. En conséquence de quoi les psycho-acousticiens en ont déduit qu'il existait en voix criée une adaptation entre phonation et audition, et que malgré un certain nombre de déformations acoustiques qui provoquent un appauvrissement de l'information transmise à l'auditeur, c'est la concentration de l'énergie spectrale dans la zone la plus sensible de l'oreille qui caractérise la voix criée. Dans le cas présent, il convient de préciser qu'à cette concentration d'énergie spectrale vient s'ajouter l'influence de l'augmentation de la durée, paramètre qui joue conjointement un rôle primaire dans la perception des voyelles en milieu bruité.



m a j w a j t r i j a k a # k i n ɔ # ʔ a k i :

Figure 4

*Spectrogramme de la séquence [majwajtrijaka#kinɔ#ʔaki:] produite par le chef du Haka (T.R.)*

#### 4.7. Les propriétés spectrales des consonnes

Dans les sections 4.2 et 4.3, nous avons insisté sur les différences entre les voyelles et les consonnes tant du point de vue de leur durée que de leur intensité. En bref, alors que les voyelles, porteuses de l'énergie sonore nécessaire à la perception du message verbal, sont favorisées et sur-articulées, les consonnes dont la fonction est uniquement de 'hacher' le flux sonore, sont en revanche négligées et sous-articulées. Sur le plan acoustique, cette sous-articulation se manifeste par la faiblesse, si ce n'est l'absence d'indices acoustiques robustes.

Cette faiblesse concerne les phases de transitions des consonnes occlusives qui sont pratiquement toujours escamotées. En effet, on ne retrouve que rarement, si ce n'est jamais, les transitions telles qu'elles apparaissent sur les spectrogrammes de parole, que ce soit en débit normal ou ralenti. Le F1 est la plupart du temps tronqué (cut-back) (Delattre, 1969), la montée de la transition de F2, indice généralement robuste en contexte dento-alvéolaire [tɛ,tã,tɑ,tɪ] est peu ou pas visible, de même que le rapprochement (connu sous le vocable de pince vélaire) des transitions de F2 et de F3 en contexte [ka, ki, kɔ]. En ce qui concerne les explosions, on note une baisse sensible de l'énergie, ce qui accrédite les remarques de Junqua '*...the burst strength decreases for all the plosives*' (Junqua, p. 514 (1992) et de Bond *et al.* (1988) qui constatent une absence de détente pour des consonnes finales produites dans des conditions de bruit rose. Compte tenu de la présence d'un bruit de fond, la visualisation des explosions 'naturelles' est particulièrement difficile sur les spectrogrammes, d'autant plus qu'elles sont parfois synchronisées avec les bruits transitoires des claquements de mains, explosions 'artificielles' de forte intensité, qui viennent compenser l'intensité faible des explosions 'naturelles', par exemple celles de [k] dans ['kɑ matɛ] ou celles de [p] dans [ʔauʔpa] (cf. section 2.3.2).

En ce qui concerne les consonnes nasales, l'élévation de la fréquence fondamentale entraîne un déplacement de l'énergie spectrale des fréquences basses vers les fréquences moyennes, et qui est semblable à celui qui a été constaté par Castellanos *et al.* (1996) pour les consonnes nasales de l'espagnol. Il semble en être de même à propos de la langue maorie, puisque le premier formant de nasalité FN1 de [m] et de [n] qui se situe habituellement aux environs de 250-300 Hz en parole normale se trouve décalé vers 400 Hz ou plus. Dans les fréquences moyennes, on relève la présence de deux formants FN2 et FN3 à 800 et 1250 Hz, qui se situent dans la continuité des formants d'oralité de la voyelle contiguë [a], à cause d'une coarticulation de type anticipatoire. Comme cette anticipation de la nasalité a été suffisamment forte pour être perçue des auditeurs, les voyelles de la syllabe initiale des mots 'ringaringa' [rĩŋarĩŋa] et 'tangata' [tãŋata] ont été transcrites comme étant nasalisées dans les deux cas.

Enfin, si les phases de transition des consonnes occlusives sont pratiquement inexistantes, il en est de même des phases de friction des consonnes fricatives. En effet, l'écoulement turbulent de l'air phonatoire se fait à travers une constriction dont la surface est nettement supérieure à la norme, ne serait-ce qu'en raison de l'augmentation de la pression sous-glottique. Comme la durée de la consonne est brève (50 ms pour [f] dans 'whiti' [fiti]) et que l'intensité du bruit est particulièrement faible, l'unique consonne fricative du *Haka* sera en règle générale peu ou mal perçue, d'autant plus que les transitions qui jouent un rôle majeur dans la perception des labio-dentales [f,v] ou des spirantes bilabiales [θ,β] (Harris,1956), ne jouent pas le rôle perceptif compensatoire qu'elles jouent habituellement dans des conditions d'écoute normale.

## 5. Les propriétés prosodiques

### 5.1. L'organisation temporelle

Dans la section 2.2, il a été indiqué que le message verbal du *Haka* pouvait se décomposer en deux parties, la harangue et le *Ka mate* proprement dit. Dans les lignes qui suivent, nous allons voir que l'un comme l'autre sont foncièrement différents tant du point de leur durée que de leur structure rythmique.

La durée de la harangue est relativement courte (11.57 sec.) ; elle comprend trois séquences composées de mots simples bi-syllabiques, par exemple *ringaringa* (main), *waewae* (pied), *kino* (mauvais), *boki* (également). Ces trois séquences sont de durée variable : la première dure à peu près deux secondes (1.967 ms) et est entrecoupée de pauses, la deuxième est sensiblement plus longue (6.548 ms) et la troisième est d'une durée intermédiaire entre les deux premières, soit 4.123 ms. Ces différences de durée montrent que l'organisation temporelle de la harangue n'est pas cohérente, et que de ce fait, elle ne se prête pas à une décomposition en vers de durée homogène.

Tel n'est pas le cas du *Ka mate* dont la durée globale (17.18 sec.) est nettement plus longue que celle de la harangue ; de plus, il s'en distingue par une structure de six vers d'une durée relativement égale : 1<sup>er</sup> vers 2920 ms , 2<sup>e</sup> vers 2855 ms, 3<sup>e</sup> vers 2822 ms, 4<sup>e</sup> vers 3205 ms, 5<sup>e</sup> vers 2555 ms, 6<sup>e</sup> vers 2828 ms (si l'on ne tient pas compte du cri final). En effet, on constate que la durée des vers 1, 2, 3, 6 est voisine, et que seuls les vers 4 et 5 sont d'une durée respectivement supérieure et inférieure à la moyenne (2.863 sec.). Comme les différences de durée entre les vers sont relativement minimales et n'excèdent pas 10%, l'auditeur a plutôt une impression d'homogénéité sur le plan temporel.

## 5.2. La structuration accentuelle et rythmique

Dans l'introduction, nous avons fait allusion à la musique du fait que, selon le terme employé dans la langue anglaise, il est fait référence au *Haka* en tant que chant. Si l'on se fonde sur l'unité de durée (mesurée) telle qu'elle a été constatée entre les vers, cette analogie avec la musique nous apparaît comme étant justifiée. En effet, au même titre que les lignes musicales sont divisées en mesures avec des notes (do, ré, mi, etc.), des silences, des soupirs, la somme de ces valeurs étant égale en durée pour toutes les mesures, les vers du *Haka* comprennent des groupes de structure uniforme, par exemple groupes trisyllabiques comme *Kamate*, *Teneite*, *Tangata*, *naitiki*, *mawaka* etc.) dont la durée est à peu près égale, ce qui accrédite l'idée selon laquelle le *Haka* est soumis à une stricte organisation temporelle, comme c'est le cas dans le domaine musical.

Afin de vérifier si cette relation avec la musique était effectivement justifiée, et de déterminer dans quelle mesure certains traits prosodiques pouvaient être perçus, un test de perception a été organisé.

### 5.2.1. Le test de perception

Selon un ordre chronologique pré-établi, il a été décidé que la tâche des sujets porterait en premier lieu sur la perception des proéminences accentuelles, ensuite sur la segmentation de chaque vers en ses différents groupes rythmiques, et enfin sur la notation des allongements vocaliques, des voyelles brèves et des pauses. Compte tenu de la difficulté de la tâche due à la mauvaise qualité d'un signal de parole produit en milieu bruité et, difficulté supplémentaire, dans une langue inconnue, nous avons eu recours à des sujets 'phonétiquement' entraînés. Ceux-ci ont été choisis parmi nos collègues du Laboratoire Parole et Langage d'Aix-en-Provence, linguistes chevronnés qui bénéficiaient pour la plupart d'une solide expérience dans l'étude des phénomènes prosodiques des langues du monde. Douze sujets ont subi le test individuellement, au cours duquel le signal sonore leur était transmis par l'intermédiaire d'écouteurs. Après un temps d'accoutumance de durée variable, le sujet demandait à l'auteur de lui faire écouter chacun des six vers du *Kamate* en sortie d'un ordinateur PC. À l'aide de divers symboles, le sujet devait marquer d'après une transcription large du *Haka*, les proéminences accentuelles ' , les groupes rythmiques | | , les allongements vocaliques \_\_ , les voyelles brèves \_ et enfin les pauses #. Les réponses des sujets ont été regroupées, analysées et une synthèse des résultats a été présentée sous la forme d'une analyse métrique.

## 5.3. L'accent

Selon l'opinion exprimée par Fraisse (1974), la perception de certaines structures prosodiques comme l'accent ou le rythme est soumise à des lois universelles. Les résultats du test confirment cette opinion, puisque nos sujets de langue française n'ont pas éprouvé de difficultés à percevoir les



proéminences accentuelles dans une langue qui leur était *a priori* totalement étrangère. Cette relative facilité s'explique du fait que l'accent de mot se situe le plus souvent en maori sur la syllabe initiale, ce qui facilite en soi la tâche de l'auditeur néophyte. De plus, comme le *Haka* se singularise par la fréquence de récurrence du marqueur de focalisation 'Ka', lequel est naturellement accentogène, la perception de l'accent s'en est trouvée facilitée. En règle générale, si les réponses des sujets témoignent d'une grande cohérence dans leur ensemble, il n'en reste pas moins que la perception de l'accent est inégale selon les vers. Si elle est maximale dans les trois premiers vers, où les syllabes 'ka, ka, kɔ (vers 1 et 2) et 'tɛ, 'tā, 'u : (vers 3) sont perçues comme accentuées dans 100 % des cas, il en va différemment des vers suivants et notamment du vers 4, où la perception de l'accent est nettement inférieure (70% seulement). Ces différences s'expliquent du fait que, contrairement aux vers précédents, le vers 4 est composé de deux groupes trisyllabiques [naitiki], [mawaka] dont la consonne initiale n'est pas une occlusive sourde [/k/] ou [/t/] en position forte, mais au contraire une consonne nasale [n] ou [m], les consonnes occlusives se trouvant dans des syllabes médianes ou finales de groupe, c'est-à-dire en position faible. On peut formuler l'hypothèse que ces différences de perception sont dues au fait que la consonne initiale de ces deux groupes est une consonne nasale [+ continu] et non une consonne occlusive [-continu]. Alors que la consonne nasale est en état de coarticulation avancée avec la voyelle contiguë [a], la consonne occlusive est sous-articulée comme en témoignent la faiblesse des explosions, ainsi que la brièveté des tenues consonantiques, cf. les syllabes 'Ka, 'kɔ, 'tɛ, 'tā. Ainsi, plusieurs sujets ont pu être sensibles à ces différences de continuité/discontinuité dans le flux sonore, et il est permis de penser que la perception de l'accent a été facilitée par la présence de ces ruptures spectrales qui hachent le continuum. Toutefois, si la perception de l'accent est relativement faible (70%) au vers 4, par contre elle est forte (85%) aux vers 5 et 6 où l'on retrouve un contexte identique à celui des trois premiers vers. En effet, les consonnes occlusives se trouvent en position forte de syllabe dans les deux groupes bi-syllabiques [ʔau'pa] du vers 5, ainsi que dans les groupes tri-syllabiques [ʔau'pane] ['kupane] du vers 6, ces syllabes ayant été perçues comme accentuées dans 90 à 95% des cas.

Sur le plan acoustique, il a été montré que, en raison de l'environnement bruité qui provoque une élévation naturelle de la voix du locuteur, les voyelles du *Haka* étaient caractérisées par une montée abrupte de la fréquence fondamentale et de l'intensité. Le fait que la F0 de [a] se situe dès l'attaque à sa fréquence maximale élimine, tout au moins limite fortement, la possibilité d'une montée mélodique ultérieure. Compte tenu des écarts minimes de F0 entre la syllabe initiale et les syllabes subséquentes, cf. les groupes trisyllabiques 'kamate, 'tɛjnetɛ, 'tangata etc., il semble que ni l'élévation de la fréquence fondamentale, ni celle du niveau d'intensité ne sauraient être considérés comme les paramètres acoustiques responsables de la perception de l'accent. Ce fait est illustré sur

les figures 1 et 2, où la syllabe initiale accentuée ne se distingue des syllabes inaccentuées subséquentes par aucune variation notable de ces deux paramètres. Cette remarque s'applique également au paramètre 'durée', vu que les mesures effectuées sur les tenues des voyelles accentuées n'ont révélé aucun allongement par rapport à leurs homologues inaccentuées. En conséquence, s'il est établi que ces variations paramétriques sont trop fines pour être perçues, on peut en déduire qu'un autre trait prosodique, en l'occurrence le rythme, joue un rôle essentiel dans la perception des proéminences accentuelles.

#### 5.4. Le rythme

Dans la troisième et dernière partie de la harangue, la séquence [ʔakinonehoki] reprise par le chœur des ABs, et qui fait suite à la même séquence produite par le chef, est assurément LE moment-clé du *Haka* dans la mesure où elle constitue une transition entre la harangue qui ne jouit pas d'une structuration rythmique véritable, et le *Ka mate* qui va être produit dès lors comme un message rythmé, scandé et cadencé. À la vérité, ce sont les claquements des mains qui vont imprimer au *Haka* son rythme de base, rythme qui sera conservé jusqu'à la fin, en dépit du fait que les claquements nettement perçus aux deux premiers vers, ne le sont plus aux vers 3 et 4, mais le sont à nouveau aux vers 5 et 6 et sont visibles sous la forme d'un agglomérat de stries verticales (figure 5) qui ne sont pas sans rappeler des explosions d'occlusives sourdes. Toutefois, comme le groupe des All Blacks est composé de vingt-deux joueurs, les claquements ne sont pas simultanés, et s'étalent dans le temps sur une durée de 100 à 150 ms. Malgré ce manque de synchronisation, un agglomérat de claquements constitue un ensemble (relativement) compact, et chacun d'entre eux est séparé du suivant par un espace temporel de 700 à 800 ms. Coïncidence sans doute non fortuite, un empan temporel de 600 à 700 ms a été trouvé comme étant celui qui s'applique au domaine musical. Cet empan temporel qui correspond en fait à 3/4 de seconde, peut être considéré comme l'écart référence-type, tel qu'il résulte d'expériences de psycho-acoustique sur l'organisation rythmique et la synchronisation sensori-motrice (Fraisse, 1974). Partant du postulat selon lequel un intervalle de 700-800 ms correspond globalement à un rythme cardiaque moyen, c'est-à-dire à des battements du pouls de 70 à 80 coups/minute, on a déterminé que cet intervalle temporel était le plus spontané dans de nombreuses activités sensorielles ou motrices. Parmi celles-ci, on mentionnera la durée d'association entre deux perceptions auditives (Wundt, 1911), la durée du pas du marcheur, la cadence du rameur, la cadence de frappe des doigts ou encore la synchronisation entre frappe et son qui se trouve être optimale pour des cadences de 600 à 800 ms, cf. la synthèse de ces expériences (Fraisse, *op. cit.*). En se fondant sur les résultats d'une expérience de 'finger tapping', Allen (1975) a émis l'hypothèse de l'existence d'une horloge neurale,

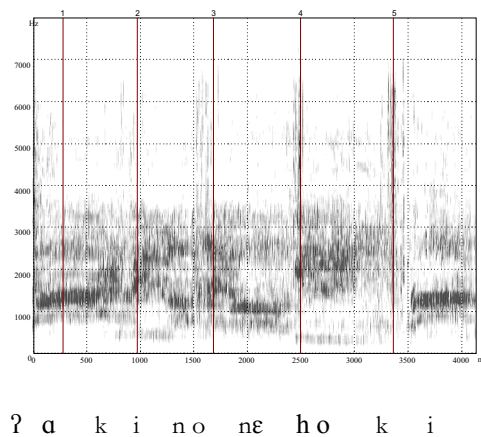
responsable selon lui de la perception du rythme dans la parole, (*...a neural clock which controls the output of both speech and other rhythmic activity*) (Allen, *op. cit.*, p. 84). Confirmant ainsi la relation qui semble exister entre rythme gestuel et parole, il a été établi que les marches, les musiques ou les danses produisent une grande activité motrice synchronisée avec les temps forts, et que les marcheurs, chanteurs ou danseurs accompagnent spontanément les sons ou temps forts<sup>9</sup> de frappes fortes. Sur ce plan, il apparaît que le *Haka* constitue assurément un exemple privilégié de synchronisation entre la voix et le geste.

#### 5.4.1. Rythme gestuel et rythme musical

Si l'on se fonde sur sa structuration rythmique, il semble donc que le *Haka* puisse être apparenté à une composition musicale. Cette analogie à la musique apparaît comme étant vérifiée, si l'on tient compte du fait que le rythme est non seulement marqué par les frappes des mains, mais encore par celles du pied, frappes qui seraient certainement audibles, si elles n'étaient étouffées par la nature spongieuse du sol, c'est-à-dire la pelouse. Ainsi, à la différence de simples récitants qui lisent un texte ou un poème, les locuteurs/joueurs peuvent être comparés à des musiciens ou à des chanteurs qui marquent la récurrence des ictus ou en d'autres termes des temps forts, par une frappe régulière du pied. Selon la définition de H. Morier, l'ictus est décrit comme le *coup frappé qui souligne le rythme*, (Morier, 1961, p. 511). D'après la terminologie grecque, l'ictus (qui existe en musique comme en poésie), est synonyme de temps fort (*'thesis'*) et s'oppose à *'l'arsis'* qui correspond pour sa part au temps faible. Dans l'interprétation d'un morceau de musique, le chef d'orchestre bat la mesure et sa main se lève entre les ictus ; à son sommet, ce temps correspond à *'l'arsis'*, alors qu'au terme de la phase descendante, le mouvement de la main correspond au *'thesis'*. Dans l'exécution du *Haka*, l'ictus (*'thesis'*) correspond à un mouvement abaisseur des bras et à un 'posé' du pied, alors que *'l'arsis'* correspond pour sa part à un mouvement élévateur des bras, et à un 'levé' du pied (ou plus exactement du talon). En se fondant sur la mesure des intervalles de 750 ms entre les temps forts ou ictus, on en déduit que le tempo moyen du *Haka* est de 75 battements/minute, ce qui correspond en langage musical à un rythme *'adagio andante'*.

---

<sup>9</sup> Le temps fort est l'équivalent musical des syllabes accentuées dans la parole.



**Figure 5**

*Spectrogramme de la séquence [ʔakinonehoki] produite par le cœur des ABs illustrant la segmentation du continuum sonore en quatre groupes rythmiques de durée équivalente*

#### 5.4.2. Rythme gestuel et rythme verbal

Après que le rythme de base ait été donné dans la dernière séquence de la harangue, le premier vers du *Ka mate* illustre la symbiose qui existe entre le rythme gestuel et le rythme vocal. En effet, se trouvent coordonnés deux types d'*ictus* : un premier *ictus* 'bruité' qui correspond à des claquements, et un deuxième '*ictus* 'vocal' ou 'mélodique' destiné à mettre en relief une syllabe donnée. À ce propos, on mesure à quel point la perception des proéminences accentuelles est dictée par des contraintes rythmiques. En fonction du découpage du continuum sonore en groupes rythmiques, trois syllabes sont mises en relief à intervalles réguliers dans chaque vers, et de ce fait sont perçues comme accentuées. Cela signifie que, même en l'absence de claquements (c'est le cas notamment aux vers 3 et 4), le sujet dont le système auditif s'est accoutumé dans les deux premiers vers à la récurrence de trois temps forts successifs, continue à percevoir trois proéminences accentuelles, en dépit du fait qu'il n'existe aucune variation acoustique majeure entre la syllabe perçue comme accentuée et les syllabes inaccentuées contiguës. Ainsi, on peut supposer que la perception de la proéminence accentuelle ne se fait pas sur une base acoustique, c'est-à-dire n'est pas fonction de l'existence de variations paramétriques conséquentes, mais plutôt sur la base de contraintes rythmiques, lesquelles ont pour effet de conditionner l'auditeur à percevoir l'accent à intervalles réguliers en fonction de la récurrence de traits para-linguistiques qui viennent se

superposer et s'ajouter au message verbal. Sur la figure 6, on remarquera la simultanéité entre l'ictus 'bruité' (claquement) et l'ictus 'mélodique' (proéminence accentuelle), ce qui atteste de la réalité de la relation entre événements gestuels et vocaux sur le plan de la structuration rythmique. En résumé, on peut dire que nous avons affaire sur le plan vocal à un accent rythmique qui coïncide sur le plan gestuel avec le frappé des mains et du pied.

### 5.5. Rythme et analyse métrique

∇(475)	∇(1272)	∇(1970)	2920 ms
ʔa:  'kamate	'kamate	'kɔ:	
—	—	—	
∇(460)	∇(1214)	∇(1952)	2855 ms
ʔa: 'kamate	'kamate	'kɔ:	
—	—	—	
Δ (468)	Δ (1194)	Δ (2005)	2822 ms
ʔa: 'tejnɛtɛ	'tāŋata	# ʔu:  ʔu:	
—	—	—	
		Δ (2125)	3205 ms
na:  naitiki	mawaka	'fiti  tea	
—	—	—	
∇(451)	∇(1913)		2555 ms
ʔau  'pa	#  ʔau  'pa		
—	—	—	
∇(450)	∇(1101)	∇(1917)	3228 ms
ʔau  'pane	'kupane	'fiti   tea  #  h i:	
—	—	—	

**Figure 6**

*Analyse métrique des six vers du Ka mate. Le symbole ∇ se rapporte aux temps forts réels, le symbole Δ aux temps forts virtuels, le symbole ' à l'accent, les traits verticaux | | aux groupes rythmiques, le trait horizontal — aux allongements vocaliques, le symbole ∪ aux voyelles brèves, et le symbole # aux pauses*

S'il est une impression auditive unanime partagée par les sujets, c'est bien celle d'une régularité rythmique ou d'un mouvement continu et régulier, qui s'annule dans sa constance et produit une impression statique (Morie, *op. cit.*, p. 933). Cette impression s'explique en fonction de la durée homogène qui caractérise aussi bien les voyelles tenues, les syllabes ou encore les groupes rythmiques à l'intérieur de chaque vers. En se fondant sur les mesures acoustiques, une analyse métrique des six vers du *Ka mate* a été faite.

### 5.5.1. Les temps forts

Par temps fort, nous entendons l'instant T où se trouvent synchronisées les deux composantes vocale et gestuelle. En d'autres termes, cela signifie que le temps fort correspond à un instant où la production d'une syllabe est accompagnée par des claquements qui vont contribuer d'une part à la mettre en relief sur le plan perceptif, et d'autre part à découper le continuum sonore en groupes rythmiques. À ce propos, nous distinguerons deux types de temps forts.

- les temps forts 'réels' sont caractérisés par la production d'une syllabe (régulièrement perçue comme accentuée) parce qu'elle se trouve être synchronisée avec des claquements audibles résultant du contact entre la paume des mains avec différentes parties du corps (cuisses, poitrine, avant-bras opposé), battements qui vont avoir pour effet de cadencer le message verbal. L'exemple type de synchronisation se situe aux vers 1 et 2, où les trois syllabes, 'Ka, 'Ka, 'kɔ sont ponctuées de battements 'réels'. Il en est de même des syllabes 'pa (vers 5) et 'pa, 'ku, 'fi (vers 6) quoique les battements résultant d'une double frappe de la main sur l'avant-bras opposé, soient moins audibles que pour les deux premiers vers.
- les temps forts 'virtuels' diffèrent des temps fort 'réels', en ce sens que la production de la syllabe n'est pas synchronisée avec un claquement 'audible', étant donné qu'il n'existe pas de contact physique entre les mains et les autres parties du corps. Les gestes consistent soit en un mouvement alternatif des bras d'avant en arrière sur un axe horizontal, soit en un mouvement du bras droit sur un axe plus ou moins vertical, les doigts étant pointés vers le sol, sans que la main droite n'entre en contact avec le bras gauche. Ces gestes sont particuliers aux vers 3 et 4, au cours desquels les syllabes accentuées 'ɛ, 'ā, 'u, 'fi ne sont pas synchronisées avec un bruit audible. Toutefois, l'importance de ces gestes ne saurait être sous-estimée dans la mesure où ils accompagnent et perpétuent le rythme verbal. En effet, d'après l'observation des films, il a été établi que la durée des séquences [ɛjnetɛ], [āŋata](vers 3) ou [naitiki] [mawaka] (vers 4) était corrélée à la durée du mouvement de protraction/rétraction des bras sur un axe horizontal ; cette durée de l'ordre de 700 à 800 ms qui correspond à l'espace temporel entre les claquements des mains aux vers 1 et 2, constitue bien la constante temporelle TYPE du Haka. L'examen de la figure 6 révèle que la récurrence de ces temps forts 'réels' ou 'virtuels' répond à une structuration rythmique rigoureuse. En effet, on constate que dans la plupart des vers, les temps forts se situent aux environs de T1 = 450 ms, T2 = 1200 ms, T3 = 2000 ms, ce qui signifie que, exception faite du vers 5 qui ne comprend que deux temps forts, les trois temps forts coïncident à la mise en relief d'une syllabe à des intervalles RELATIVEMENT égaux. Cette cadence confirme

la pertinence de la remarque de Bolinger (1985), selon laquelle le locuteur impose un équilibre rythmique à sa production vocale. Sans que cela lui soit imposé par la structure syntaxique et grammaticale de l'énoncé, il contrôle et choisit l'occurrence des syllabes proéminentes afin de promouvoir le rythme adéquat (voir également Astésano, 2000).

#### 5.5.2. Les écarts inter-accentuels

Ainsi, il semble que l'accent soit perçu le plus souvent en rapport avec un temps fort ponctué par une frappe des mains ou par un mouvement des bras, gestes qui accompagnent le message verbal et lui donnent sa rythmicité. Conséquence de cette récurrence de temps forts, l'auditeur pourrait avoir l'impression d'un isochronisme entre les syllabes accentuées, c'est-à-dire d'une préservation d'une stricte régularité temporelle. Toutefois, se pose la question de la réalité de cette régularité. En termes de données quantitatives, ces écarts inter-accentuels sont-ils constants ou au contraire sujets à variation ?

Les mesures de durée montrent qu'il n'existe pas de traces acoustiques objectives de la rythmicité. La seule constante réside en la récurrence d'un premier temps fort associé à la perception de l'accent, et qui est invariablement situé à  $T_1 = 450$  ms. En revanche, il n'existe pas de stricte régularité temporelle entre les différentes syllabes accentuées dans chacun des six vers du *Haka*. Si la durée des écarts inter-accentuels avoisine 750 ms dans les trois premiers vers, soit une valeur intermédiaire entre 700 et 800 ms, cf. l'intervalle référence (section 5.4.), ces mêmes écarts inter-accentuels sont variables dans les trois vers suivants, notamment au vers 6, où l'écart entre la 1<sup>e</sup> et la 2<sup>e</sup> voyelle accentuée n'est que de 651 ms, alors qu'il est de 816 ms entre la 2<sup>e</sup> et la 3<sup>e</sup> voyelle accentuée.

En conclusion, nous dirons que l'isochronisme des écarts inter-accentuels est d'ordre relatif et non pas absolu. Cette assertion va dans le sens de celle de Fant (1991) pour lequel la rythmicité accentuelle n'est pas due à un isochronisme physique rigoureux, mais plutôt à une dominance perceptive de syllabes fortes dites 'heavy', dont la succession est ressentie par l'auditeur comme quasi périodique. Son opinion conforte celle de Lehiste (1973) selon laquelle l'isochronie n'a de réalité que dans la perception, l'audition primant la production. Enfin, Ramus *et al.* (1999) confirment les opinions exprimées par ces deux derniers auteurs, quand ils laissent entendre que les intervalles accentuels ne semblent pas avoir de durée constante, étant donné que leur durée intrinsèque dépend aussi bien du type de syllabes que de la position de l'intervalle dans l'énoncé.

#### 5.5.3. Les groupes rythmiques

Les résultats du test de perception montrent que chacun des six vers est susceptible d'être décomposé en plusieurs groupes rythmiques. En règle générale, ces groupes sont au nombre de

quatre (vers 1, 2, 3, 5), seuls les vers 4 et 6 en comprennent cinq en raison d'une durée plus longue (en particulier le vers 6 à cause du cri final). Si la durée des groupes est plus ou moins constante en raison des claquements de mains ou des mouvements des bras qui rythment la cadence verbale, la structure syllabique des groupes diffère selon qu'ils se trouvent en position initiale, médiane ou finale de vers.

En position initiale, le groupe rythmique est invariablement constitué d'un monosyllabe. D'une durée moyenne de 450 ms, ce monosyllabe peut être assimilé à une *anacrouse*, terme habituellement employé en poésie ou en musique par les grammairiens de l'Antiquité, et repris à son compte par Jassem (1999) dans son modèle de prédiction de la durée des unités rythmiques en anglais britannique.

- dans le domaine de la poésie et plus particulièrement dans le vers libre, mais parfois aussi dans la prose, l'*anacrouse* se rapporte à la syllabe qui sert de prélude à la cadence métrique. Selon Morier (*op. cit.*), la syllabe en *anacrouse* est constituée d'une conjonction de coordination (mais, donc, ni, etc.), quelquefois d'une onomatopée (chut !) ou le plus souvent d'une exclamation du type *Ab !*. Or, il se trouve que c'est une exclamation de ce type qui se rencontre à l'initiale de vers sous la forme de la voyelle glottalisée [ʔa], la montée tonale étant ponctuée sur le plan gestuel par une élévation simultanée des bras. Si dans un vers classique, l'*anacrouse* (l'exclamation) précède en général le groupe polysyllabique, on constate qu'il en est de même dans le *Haka*, où l'*anacrouse* (la voyelle initiale) produite pendant l'*arsis* (le temps faible) précède le groupe polysyllabique, cf. [ʔa 'kamate] dont nous rappelons que la syllabe initiale est porteuse de l'ictus mélodique et correspond au *thesis* (temps fort) marqué par les battements de mains.
- dans le domaine de la musique, l'analogie se trouve renforcée du fait que l'*anacrouse* désigne également la ou les notes initiales d'un rythme qui mènent au premier temps fort. C'est ce cas de figure qui se rencontre à l'attaque de vers, où la voyelle [a], en raison de sa fréquence d'attaque élevée (>400 Hz) et de sa durée (>450 ms), fait figure de note tenue qui en l'occurrence pourrait être une ronde. On ajoutera que, si l'*anacrouse* correspond en musique à la note ou aux notes qui précèdent la première barre de mesure, l'analogie avec la voyelle initiale [a] du *Haka* s'impose, du fait que la barre de mesure musicale apparaît sur le spectrogramme (figure 5) sous la forme de l'agglomérat de stries verticales que nous avons décrit.

En position médiane de vers, les deuxièmes et troisièmes groupes sont constitués de polysyllabes, par exemple *kamate*, *teneite*, *tangata*, *mawaka* etc. Ces polysyllabes sont composés de syllabes brèves dont la durée moyenne est de 248.80 ms. Toutefois, les données acoustiques du tableau 3 indiquent qu'il existe des variations temporelles entre lesdites syllabes ; par exemple dans le groupe



*kamate* du vers 2 apparaissent des différences de l'ordre de 66 ms entre la première et la deuxième syllabe, soit un écart de 25 %. Ce fait implique que, à l'instar de l'isochronie accentuelle qui n'existe que sur le plan de l'impression auditive et non pas sur celui de la réalité acoustique, le phénomène d'isosyllabité, c'est-à-dire une durée égale entre les syllabes du même groupe rythmique, ne s'avère pas vérifié. Si les trois premiers vers présentent en position interne une structure syllabique uniforme, c'est-à-dire qu'ils sont composés de deux groupes rythmiques où les syllabes sont organisées en rhèmes ou mots phonétiques équilibrés sur le plan temporel, il n'en est pas de même dans les trois vers suivants. Les raisons de cette disparité doivent être recherchées dans les différences de durée entre les vers. En effet, nous avons remarqué que les vers 4 et 6 se distinguent des autres vers par une durée plus longue, laquelle est due à la présence d'un quatrième groupe rythmique [ʼfitʃ], d'où un enchaînement de huit syllabes brèves au lieu de six syllabes dans les vers 1, 2, 3. De son côté, le vers 5 se distingue par sa brièveté, ainsi que par une structure syllabique différente de celle des autres vers (production de bi-syllabes au lieu de trisyllabes) et par l'insertion d'une pause (*cf.* section suivante).

Enfin, on remarquera que, en position finale de vers, le groupe rythmique apparaît sous la forme d'un monosyllabe. À la différence de la position initiale, où ledit monosyllabe est constitué d'une voyelle tenue isolée, le groupe rythmique terminal comprend soit une syllabe ouverte [kɔ] (vers 1 et 2), soit une voyelle tenue [u] (vers 3) ou [a] (vers 5), soit une syllabe avec des voyelles enchaînées [tea] (vers 4 et 6). Quelle que soit la forme sous laquelle il apparaît, ce groupe rythmique terminal diffère des groupes médians trisyllabiques, du fait que sa durée est la plupart du temps supérieure à 800 ms. Cette voyelle terminale, qui constitue en quelque sorte le corollaire de la voyelle initiale, participe à l'équilibre temporel du vers, et fait office de note finale qui conclue habituellement le couplet ou le refrain d'une chanson.

En résumé, l'analyse métrique des vers qui composent le *Ka mate* laisse apparaître une structure rythmique fondée sur une succession de voyelles longues et de voyelles brèves. La séquence type peut être représentée graphiquement comme suit — ◡ ◡ ◡, c'est-à-dire une voyelle longue suivie de trois brèves, la première d'entre elles étant porteuse de l'accent. Cela signifie que la structure métrique du *Haka* est de type trochaïque avec une tête métrique à gauche. Par ailleurs, le fait qu'il existe une régularité de pieds accentuels, ou en d'autres termes une égalité temporelle entre les accents, semble indiquer que le maori est, à l'instar de la plupart des langues polynésiennes, une langue à rythmicité accentuelle.

## 5.6. Les pauses

Dans l'activité langagière, on distingue deux types de pauses ; d'une part les pauses silencieuses grammaticales qui coïncident avec des frontières de groupes syntaxiques ou prosodiques, d'autre part, les pauses d'hésitation où le choix des mots est fonction de la connaissance qu'a le locuteur du thème ou du sujet (Duez, 1982).

Dans la situation de communication verbale insolite où se trouvent les exécutants du *Haka*, l'insertion de pauses répond à un tout autre besoin, et participe à ce que Fónagy appelle *la stratégie dramatique des silences*. Cette dernière expression est particulièrement appropriée dans cette exhortation au combat qu'est le *Haka*, puisque les silences « *Ces brèches dans le discours* » (Redelücken) ou encore ces instants où « *...il s'y passe quelque chose, même si rien n'est dit ; il s'y crée des états de tension psychologique et physique qui se transmettent du locuteur à l'interlocuteur...* » (p. 147, Otto von Essen (1957 ; cité par Fónagy, *op. cit.*, p. 256), accentuent l'intensité dramatique de l'affrontement qui va s'ensuivre. Dans le cas qui nous préoccupe, l'insertion de pauses est également destinée à favoriser la respiration et à maintenir l'équilibre temporel et rythmique entre les vers.

### 5.6.1. La respiration

Cette fonction respiratoire est mise en évidence pendant les pauses qui précèdent la harangue. Sur ce plan, l'observation des films vidéo permet de constater que le chef et ses équipiers se livrent à de profondes inspirations suivies d'expirations avant l'exécution du *Haka*. Ces exercices respiratoires, qui ne sont pas sans rappeler ceux des plongeurs en apnée avant l'immersion, ont un double but : d'une part réguler le rythme cardiaque afin de maîtriser les émotions et la tension nerveuse, d'autre part accumuler un maximum d'air pulmonaire avant la production orale. Après avoir jeté un dernier coup d'œil circulaire sur ses équipiers, le chef du *Haka* se retourne brusquement vers l'équipe adverse, et produit à deux reprises la séquence [karitɛ], chacune d'entre elles étant suivie d'une pause d'environ 200 ms. Ensuite et après que ses équipiers aient répondu à la séquence [ʔikamau] sous la forme d'un long hululement, le chef marque une troisième pause (400 ms) avant d'attaquer la deuxième partie de la harangue [ʔarɪŋarɪŋa...]. Comme celle-ci dure plus de 6 secondes, trois pauses de 100 à 150 ms seront observées afin de reprendre son souffle et produire la séquence entière à un niveau constant de puissance sonore. C'est une situation identique que l'on retrouve à la fin du *Ka mate*, où l'insertion d'une pause est fonction de la durée du vers 6 (3.2 sec.). En effet, avant la fin du vers, une pause de 400 ms est observée et est mise à profit par les joueurs/locuteurs pour inspirer une ultime bouffée d'air avant de pousser leur dernier cri synchronisé avec le saut final.

On ajoutera que durant les temps morts qui précèdent l'exécution du *Haka*, la tension psychique atteint son paroxysme ; les muscles se contractent, les poings se crispent, les regards se font haineux et provocateurs, les yeux sortent de leurs orbites, la poitrine se gonfle, la pomme d'Adam se soulève et la gorge se serre, prête à éructer les imprécations les plus terribles. Toutefois, dans ce contexte de confrontation, le rôle des pauses ne se limite pas à la seule fonction de respiration ou d'expression des émotions, il est également important dans le maintien de l'équilibre rythmique.

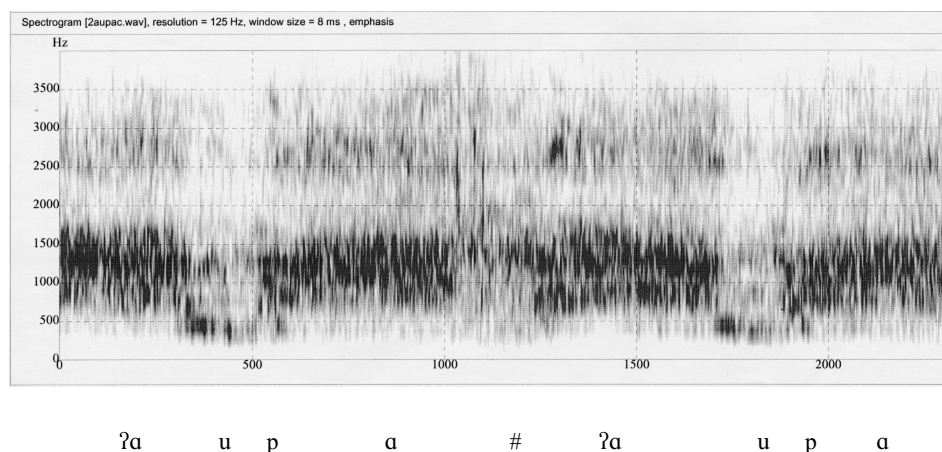
## 5.6.2. L'équilibre rythmique

### 5.6.2.1. Les pauses réelles

Le vers 5 est l'exemple type d'une pause (césure en langage poétique) observée à dessein pour préserver l'équilibre rythmique et assurer une égalité de durée entre les vers. Dans ce vers, l'insertion de la pause est due à la brièveté relative de ce dernier (2.555 ms). Comme le montre l'analyse métrique (figure 6), cette différence de durée n'est pas due au fait que ce vers comprenne moins de groupes rythmiques que les autres vers (dans un cas comme dans l'autre, on en dénombre quatre), mais plutôt au fait que deux de ces groupes rythmiques soient formés du monosyllabe | *pa* | dont la brièveté contraste avec la longueur des groupes tri-syllabiques internes, cf. | *Kamate* | *teneite* | *tangata* | *naitiki* | *mawaka* | etc.

En raison de l'insertion de la pause, le flux sonore va se trouver interrompu au milieu du vers. Cette rupture va avoir pour effet d'allonger sa durée, mais aussi de rétablir l'équilibre rythmique, lequel aurait été rompu en raison de la brièveté du vers. Selon H. Morier, ce point césurien (pc) qui se situe sur le plan sonore en dehors des accents constitue le *...moment creux du vers, le lieu central, là où fléchit la sonorité* (Morier, *op. cit.*, p. 188). On ajoutera que ce point césurien (pc) – on pourrait également utiliser le terme de temps mort – tombe entre deux temps forts marqués par les ictus accentuels et bruités qui mettent en relief les deux syllabes *pa*. Dans ce cas, la césure a pour fonction de palier l'absence du deuxième temps fort qui se situe généralement à  $T = 1200$  ms. Ainsi, se trouve respecté l'équilibre rythmique du vers où les ictus 1 et 3 sont respectivement marqués à  $T_1 = 500$  ms et  $T_3 = 2000$  ms. À titre d'illustration, on observera sur le spectrogramme de la figure 7 que d'une part, le point césurien coïncide avec le milieu temporel entre les deux ictus (en d'autres termes, les ictus sont disposés symétriquement de part et d'autre de la césure), et que d'autre part, le milieu temporel est déterminé par les bornes extérieures du vers, cf.  $pc = 1.250$  ms, c'est-à-dire un intervalle de temps intermédiaire entre 0 et 2555 ms. Notre remarque va dans le sens de celle de H. Morier qui a montré, à l'appui de mesures effectuées sur des oscillogrammes,

que ... *le milieu temporel se rencontre, à quelques centièmes de seconde près, avec le point césurriel phonétique* (Morier, *op. cit.*, p. 189).



**Figure 7**

*Spectrogramme du vers 5 ʔau ʔa # ʔau ʔa illustrant l'existence d'une pause 'réelle' de 200 ms située à environ T = 1250ms, c'est-à-dire entre les bornes extérieures du vers*

#### 5.6.2.2. Les pauses virtuelles ou pauses subjectives

Si la pause 'réelle' se manifeste sur le plan acoustique par une rupture physique du continuum sonore, il existe un autre type de pause 'virtuelle' ou pause 'subjective' qui, à la différence de la précédente, n'entraîne pas d'interruption dans ledit continuum. Cette césure ou *pseudo-césure mélodique* (Morier, *op. cit.*, p. 193) est en fait une pause perceptible, mais inexistante sur le plan acoustique, et est marquée *d'une certaine façon par la mélodie de la phrase* (Zwirner, 1937). C'est une pseudo-césure de ce type qui caractérise la voyelle terminale [u:] du vers 3. Comme le montre l'analyse métrique, cette voyelle a été perçue en deux unités successives. L'indice acoustique responsable de cette impression auditive est une chute mélodique de 60 Hz qui va de pair avec une chute d'intensité de 6 dB. Comme cette chute mélodique d'une durée de 150 ms affecte la partie médiane de la voyelle, l'auditeur a l'impression de percevoir une 'très faible pause'. En réalité, celle-ci est due à la chute concomitante de la fréquence fondamentale et de l'intensité, laquelle est suivie d'une remontée progressive de ces deux paramètres sur la partie terminale de la voyelle. Ainsi, l'auditeur a l'impression d'un 'ton flottant' parce que la voyelle (qui se trouve généralement en position finale dans un vers) est fortement allongée, *...cet allongement étant associé à la pause qui la*

*Spectrogramme de la voyelle finale [u:] (vers 3) expliquant la perception d'une pause 'virtuelle' en fonction de l'existence d'une chute mélodique sur la partie médiane de la voyelle*

Au terme de cette étude consacrée à l'analyse acoustique du signal de parole produit par le chœur des All Blacks, il a été démontré que le contexte tant psychologique que physique qui entoure l'exécution du *Haka* avait une influence prépondérante sur le style de parole adopté par les joueurs/locuteurs.

72

s'exprime sur le plan oral et gestuel dans le *Haka*, avant de s'exprimer physiquement sur le terrain. C'est ce caractère rassembleur du *Haka* que T. Randell met en exergue quand il déclare '*...the important thing is to read pride and determination in the eyes of my team-mates. The fear of my opponents does not matter to me*'. À ce premier facteur de motivation, vient s'ajouter un deuxième facteur qui a trait au climat de rivalité et de passion qui entoure souvent les matches des All Blacks. Dans une confrontation avec ces derniers, l'adversaire se trouve, toutes proportions gardées, et même si cela s'avère moins vrai aujourd'hui avec l'avènement du professionnalisme, dans la situation de l'élève vis-à-vis du maître, et de ce fait va mettre toute sa force et son énergie pour éventuellement le surpasser. Cette situation d'équipe à battre est bien connue des All Blacks qui savent qu'il leur faudra toujours faire preuve de davantage de pugnacité, d'adresse et de vitesse pour vaincre. À la différence de la plupart des autres équipes de rugby, exception faite de certaines équipes de Polynésie qui exécutent également une danse guerrière avant le match (en particulier les rugbymen des îles Tonga ou Samoa), c'est dans le *Haka*, chant guerrier UNIQUE apte à véhiculer les émotions les plus fortes, que les joueurs néo-zélandais expriment leur volonté de vaincre. À ces facteurs psychologiques qui vont influencer sur le mental des joueurs, vient s'ajouter le facteur de l'environnement physique particulier au stade, théâtre moderne du déchaînement des passions. Comme nous l'avons fait remarquer, le défi lancé à l'adversaire est souvent interprété comme une provocation de la part de la foule dont les clameurs interfèrent avec le signal vocal émis par les locuteurs/joueurs. Par réaction, ces derniers vont élever le niveau sonore de leurs voix individuelles, afin que le message soit entendu haut et fort, élévation qui va occasionner des changements importants dans la structure acoustique des sons ainsi que dans la structure prosodique de l'énoncé.

Dans ce contexte passionnel où la perception du message est assurément plus importante que sa compréhension, les All Blacks vont avoir recours à un certain nombre de procédés tant paralinguistiques que linguistiques pour que ledit message soit perçu de l'adversaire. Le recours à ces procédés va s'avérer d'autant plus indispensable que, le texte écrit ne comportant aucun terme agressif, c'est uniquement par le biais ou dirons-nous par la magie du vocal ou du gestuel, la forme primant le fonds, que les locuteurs/joueurs vont transformer un texte en apparence non violent sur le plan sémantique, en une harangue guerrière agressive et vindicative. Sur ce point, on ajoutera que les joueurs sont unanimes à reconnaître que le texte du *Ka mate* continue à avoir, ne serait-ce que pour eux-mêmes, une grande importance dans son aspect traditionnel, '*...Ka Mate is still central to the All Blacks culture*' (T. Umaga)<sup>10</sup>.

---

<sup>10</sup> '*All Blacks unveil new Haka*', 29/08/2005, <allblacks.com> in Dunedin.

Parmi les procédés para-linguistiques (non vocaux) qui concernent l'attitude corporelle et l'expression faciale, certains d'entre eux n'ont pas d'incidence directe sur la production orale, c'est le cas notamment du déhanchement, de la vibration des mains '*kakapa*', des mimiques faciales '*pukana*, *whetero*' etc. Par contre, d'autres procédés, comme les battements de mains ou du pied, sont importants dans la mesure où, synchronisés avec la voix, ils concourent à marquer le rythme de base, contribuant ainsi à mettre en relief certaines syllabes et à donner l'impression auditive d'une parole dure, scandée et martelée.

En ce qui concerne les procédés linguistiques, on constate que l'utilisation systématique de certains d'entre eux est à l'origine de la disparité entre production orale et texte écrit. Parmi les procédés les plus utilisés, on retiendra l'adjonction de la voyelle [a] à l'attaque de chaque vers, mais aussi l'élision de syllabes, procédé par lequel des groupes pluri-syllabiques se trouvent souvent réduits à des monosyllabes, ce fait étant dû aussi bien à des contraintes articulatoires que rythmiques.

À plusieurs reprises, nous avons fait allusion au concept de 'dureté' qui est associé à la force mâle et virile du guerrier, mais aussi à la puissance de son organe vocal. Afin que sa voix porte le plus loin possible, le locuteur/joueur va adapter sa stratégie de communication aux conditions de bruit ambiant, et va privilégier la production de voyelles qui sont porteuses de la majeure partie de l'énergie acoustique. Cette prédilection concerne la voyelle [a] qui se trouve être à la fois la plus longue et la plus intense de tout le système vocalique. De surcroît, la sur-pression glottique favorisant la formation d'un coup de glotte pré-vocal, la glottalisation systématique de la voyelle [a] en position initiale va accentuer l'impression de dureté de la parole. Outre le fait que la voyelle la plus ouverte se trouve être également la plus fréquente sur le plan distributionnel, on note parallèlement une tendance à l'ouverture chez les autres voyelles, qu'il s'agisse des voyelles fermées [i]-[u] ou des voyelles mi-fermées [e]-[o], cette tendance s'expliquant par le fait que les voyelles soumises à l'effet Lombard sont produites en tant que variantes plus ouvertes. De plus, si les voyelles se distinguent par une ouverture accrue par rapport à la norme, elles se distinguent également par un allongement considérable des tenues. Le fait que le rapport moyen de durée entre les voyelles et les consonnes soit de 3.50 à 5.05 en faveur des premières, atteste de la réalité d'un vocalisme dominant, l'unique fonction des consonnes étant de 'hacher' le continuum sonore.

Dans l'introduction, était posée la question de savoir à quel type de voix correspondait la voix produite par le chœur des All Blacks, et dans quelle mesure ce type de voix devait être assimilé à de la voix parlée, criée ou chantée. L'analyse acoustique permet d'apporter plusieurs éléments de réponse.

En ce qui concerne le premier type de voix, c'est-à-dire la voix parlée, la réponse est claire et non ambiguë. Compte tenu des conditions d'émission du message verbal, la voix des All Blacks ne saurait être assimilée à une voix que l'on qualifierait de 'normale'. La tension psychologique a pour

conséquence une tension physiologique de l'appareil phonatoire tout entier, et en particulier du mécanisme laryngé, tension qui provoque une élévation de la fréquence fondamentale, de l'intensité globale ainsi qu'une augmentation de la durée vocalique. Cette augmentation concomitante des paramètres a pour effet que la voix des locuteurs est de toute évidence 'forcée'. À la vérité, nous avons affaire à une élocution inhabituelle ou en d'autres termes à une parole 'émotionnelle' au sens propre du terme, style de parole qui l'écarte radicalement d'un mode de production plus ou moins conventionnel.

De façon explicite, la réponse à la question concernant le type de voix inhérent au chœur des All Blacks se trouve contenue dans la métaphore utilisée par certains journalistes ou commentateurs sportifs, par laquelle le *Haka* est décrit en tant que « cri de l'âme maorie ». En effet, si l'on se place sur le plan de la réalité physique brute, on ne peut que constater, à l'appui des données quantitatives que nous avons recueillies, que la voix produite par les All Blacks, présente effectivement toutes les propriétés acoustiques de la voix criée.

En premier lieu, une fréquence d'attaque de F0 supérieure à 400 Hz, c'est-à-dire supérieure d'une octave et d'une quarte<sup>11</sup> par rapport à la fréquence moyenne de voix masculines (150 Hz), montre que les locuteurs sont à l'extrême limite de leur tessiture. Cette élévation de la F0 entraîne des variations de fréquence des formants, en particulier une élévation de la fréquence de F1 en raison d'une grande ouverture dans le conduit vocal. Conjointement, on remarque des variations de fréquence de F2, lesquelles ont pour résultat un mouvement centralisateur des fréquences cibles des voyelles, et en particulier des voyelles fermées [i] et [u]. Le resserrement du trapèze acoustique, qui implique une déviation par rapport aux configurations articulatoires canoniques, entraîne une confusion perceptive et une moins bonne intelligibilité des voyelles criées par rapport à leurs homologues parlées, ceci étant dû au fait que les locuteurs cherchent davantage à produire de l'énergie sonore qu'à transmettre une information linguistique ou sémantique. En ce qui concerne l'opposition entre les voyelles 'criées' du *Haka* et les voyelles 'parlées', on remarque une différence d'amplitude des harmoniques. Cela signifie que l'énergie n'est pas concentrée dans le premier harmonique et le F1, mais a tendance à s'étaler vers les fréquences moyennes du spectre audible, c'est-à-dire jusqu'à 4.5 KHz, où l'énergie chute brutalement à partir de cette fréquence. Enfin on ajoutera que, même si ces variations de fréquence des formants sont solidement établies, elles restent néanmoins limitées, et ne sont pas proportionnelles aux variations considérables de F0 que nous avons constatées entre ces deux types de voix.

---

<sup>11</sup> Calcul d'un intervalle en savarts :  $1000 \log_{10} 400 \text{ Hz}/150 \text{ Hz} = 2.66666647$  ;  $\log_{10} = 0.425968732 \times 1000 = 425,966$  savarts = 8.5 tons = une octave et une quarte.



En ce qui concerne l'intensité, nous avons expliqué les raisons pour lesquelles l'intensité mesurée d'après la courbe d'amplitude globale ne correspondait pas à l'énergie telle qu'elle était rayonnée aux lèvres des locuteurs/joueurs. En termes d'intensité réelle, le niveau est certainement très élevé (de 85 à 90 dB), soit un niveau très supérieur à celui d'une voix normale (environ 60-65 dB). L'augmentation du niveau sonore de la voix a pour conséquence de provoquer non seulement un écrêtage des amplitudes (les variations d'intensité ont un aspect rectangulaire), mais aussi un écrêtage fréquentiel qui donne aux voyelles un aspect de voyelles saturées. À ce sujet, le niveau de cette énergie sonore hors normes est attesté par le réflexe qu'ont la plupart des téléspectateurs ou auditeurs, et qui est de baisser spontanément le niveau de sortie du son du téléviseur ou du magnétophone dès les premières séquences verbales du *Haka*.

Enfin, en ce qui concerne la durée des voyelles, celle-ci est de loin supérieure à celle de voyelles qui auraient été produites dans des conditions de parole lue, normale ou spontanée. À titre de comparaison, on notera que la durée des voyelles brèves du *Haka* (en contexte CV) est de l'ordre de 200 ms, soit une durée voisine de la durée moyenne de voyelles accentuées en parole normale. Par contre, la durée des voyelles longues s'échelonne entre 450 et 900 ms. C'est le cas de la voyelle [ɑ:] produite à l'initiale des vers (450 ms), des voyelles finales (750 à 900 ms), et plus encore, des cris assimilés à des voyelles tenues, et dont la durée excède 1000 ms. Cet allongement inusuel des tenues vocaliques, qui a pour contre-partie un abrègement des tenues consonantiques, atteste d'une tendance à la sur-articulation des voyelles, et *a contrario* d'une tendance à la sous-articulation des consonnes, comme en témoignent la faiblesse des explosions et l'absence de transitions. Comme il a été démontré que cette tendance inverse caractérise les rapports de durée entre voyelles et consonnes en voix chantée, on ne peut s'empêcher d'établir un rapprochement entre la voix du chœur des All Blacks et ce dernier type de voix.

Ce rapprochement ne manquera pas de soulever certaines objections de la part de puristes du chant qui avanceront que, ne saurait être considérée comme un chant une harangue guerrière produite sous forme de cris gutturaux ou stridents, dénuée de toute harmonie vocale, et qui plus est, n'est caractérisée par la présence d'aucune variation mélodique susceptible d'évoquer la musique ou le chant. S'il est un fait que ces variations sont de faible amplitude, et que le contour mélodique peut être comparé, du point de vue de sa représentation graphique, à un long plateau pratiquement uniforme, on peut avancer plusieurs contre-arguments qui militent en faveur de la thèse inverse, et selon laquelle le *Haka* peut être également assimilé à un chant.

Le premier de ces arguments est relatif au fait qu'une terminologie à connotation musicale est souvent utilisée dans la littérature anglo-saxonne pour faire référence au *Haka*; parmi les termes les plus fréquemment employés, on relèvera les termes d'incantation, de 'refrain', de 'musical composition' et

surtout le terme de ‘chant’, quoique celui-ci implique dans la langue anglaise, une parole ‘scandée’ et ‘martelée’, nuance qui n’est pas présumée dans sa traduction littérale en français.

Le deuxième argument concerne le fait que certaines parties du corps, par exemple les pieds, les mains, les bras, soient comparés de façon métaphorique à des instruments dont l’exécutant fait usage comme le ferait un musicien dans un orchestre, ‘*The Haka is a composition played by many instruments. Hands, feet, legs, body, voice, tongue and eyes, all play their part in blending together...*’ (Armstrong, 1964, cité par Karetu, *op. cit.*, p. 25). Dans cet ordre d’idées, on notera que chez certaines tribus africaines (notamment les zoulous), le bruit résultant du contact de la lance avec le bouclier ponctue les cris de guerre et cadence la marche vers l’ennemi.

Le troisième argument sera protéiforme tant la référence faite à la musique est omniprésente :

- le chef fait office de chef d’orchestre, mais à la différence de ce dernier qui utilise une baguette, il utilise ses bras pour battre la mesure et marquer le rythme.
- la structure rythmique du *Haka* s’apparente à celle d’un morceau de musique divisé en mesures, en ce sens que les vers sont décomposables en groupes rythmiques de durée égale. Sur le plan de la représentation graphique, on notera l’analogie entre les barres verticales des spectrogrammes de parole et les barres de mesure des partitions musicales.
- les syllabes initiales accentuées des groupes trisyllabiques sont ponctuées de battements, et sont assimilables à des temps forts musicaux qui se situent sur le premier temps de chaque mesure. À cet égard, les deux premiers vers du *Ka mate* sont l’illustration de cette symbiose entre prééminences accentuelles et temps forts musicaux [ʔa:’kamate’kamate’ko:]. De plus, leur succession à intervalles réguliers (750 ms) fait que le rythme du *Haka* est en langage musical l’équivalent d’un *adagio andante* à 75 battements/minute.
- si l’on pousse l’analogie plus loin, on constate que si, dans le domaine musical, une note est généralement émise avant le premier temps fort, il en est de même dans le *Haka*, où la première syllabe accentuée (qui est en fait l’équivalent du temps fort) est invariablement précédée d’une voyelle tenue au début de chaque vers.
- de plus, il est permis d’établir un parallèle entre le chanteur qui émet une note de référence à une hauteur fixe (le ‘la’ du diapason) avant une vocalise ou une chanson, et le locuteur/joueur qui produit une voyelle de même hauteur, laquelle lui servira de voyelle référence à l’attaque de chaque vers.
- enfin, l’analyse acoustique a mis en évidence le fait qu’il existait un rapport de durée globale de 1 à 2 entre voyelles longues et voyelles brèves, c’est-à-dire un rapport du simple au double. En effet, les mesures acoustiques montrent que la durée de la voyelle initiale (400-450 ms) est la plupart du temps l’équivalent du double de celle des voyelles brèves (200-

230 ms) en contexte CV. Là encore, l'analogie avec la musique est patente, les voyelles pouvant être comparées à des notes ; les voyelles tenues sont le plus souvent deux fois plus longues que les voyelles brèves, comme le seraient en musique les rondes par rapport aux blanches, ou encore les blanches par rapport aux noires.

Au tout début de cette étude, nous posions en préambule la question de savoir quelle pourrait être la terminologie la plus appropriée pour qualifier le type de parole produite par les All Blacks pendant l'exécution du *Haka*. À cette question, il est difficile de répondre de façon univoque, tant il est vrai que l'on peut adopter des points de vue différents à ce sujet. Si l'on retient avant toute chose le caractère dur, puissant, agressif ou martelé de la parole, il est vraisemblable que le terme de 'voix criée scandée'<sup>12</sup> fondée sur un contrôle temporel plutôt que musical, soit le terme adéquat pour rendre compte du caractère non mélodieux d'un message dépourvu de variations tonales, et qui de ce fait n'est ni agréable, ni plaisant sur le plan perceptif.

Si l'on se situe sur le plan acoustique, et que l'on tient compte du fait que la 'voix chantée criée' (voix inculte des débutants et des amateurs) est caractérisée par un renforcement de TOUS les harmoniques, la 'voix chantée cultivée' étant caractérisée en ce qui la concerne par un renforcement d'harmoniques dans le grave et vers 4000 Hz, alors on peut considérer le *Haka* en tant qu'une 'voix chantée criée' ou encore une 'voix chantée inculte'<sup>13</sup>. À ce propos, ne dit-on pas d'un chanteur qui pousse sa voix, c'est-à-dire dont la pression subglottique est très élevée 'qu'il crie, mais qu'il ne chante pas' (Scotto et Autesserre, 1993).

Enfin, si l'on se place sur le plan métaphorique, il semble tout aussi légitime de considérer que le *Haka* en tant que moyen d'expression et d'extériorisation des émotions, est un chant, chant par lequel s'expriment la passion, la vigueur et l'identité de la race maorie selon les termes employés par Armstrong (1964) *...this complex dance is an expression of the passion, vigour and identity of the race. It is at its best, truly, a message of the soul expressed by words and posture...*(cit. par Karetu, *op. cit.*, p. 25). À la vérité, si le *Haka* est l'émanation des sentiments authentiques qui viennent du plus profond de l'âme, alors il convient de reconnaître qu'un tel message ne peut être que chanté, opinion partagée par O. Jespersen quand il déclare *Out of the full heart, the mouth sings*, (Jespersen, 1922, p. 434). Et en fait, les exemples abondent dans les textes littéraires, où les hommes se sont plu à chanter de tout temps et de tout leur cœur, les exploits guerriers et les hauts faits de leurs ancêtres., *...exploits of war, the deeds of their ancestors, any incident is turned into a song* (Jespersen, *op. cit.*, p. 434). Nul doute que dans le *Haka* sont évoquées et chantées non seulement les vertus de courage et de bravoure de la

---

<sup>12</sup> Suggestion de notre collègue D. Autesserre.

<sup>13</sup> Ces différents termes nous ont été suggérés par N. Scotto Di Carlo.

race maorie, mais aussi implicitement les exploits et les victoires des glorieux anciens qui ont partagé, sur tous les terrains du monde, la même passion pour ce jeu de rugby, et dont les présents exécutants se veulent les dignes successeurs. C'est cette passion éternelle pour le maillot noir qui anime T. Umaga, dernier capitaine en date des All Blacks quand il déclare à l'issue de son premier *Haka* exécuté en tant que chef du chœur : ‘ *We were told that the words do mean a bit, but it is how you go about it, and the feeling and passion that comes from within (that matters). Hopefully, I gave everything I had*’ (from ‘Umaga lays his ghosts to rest’, *The Dominion Post*, Nov. 29<sup>th</sup>, 2004). Et de fait, si le chant n'est que le résultat quasi automatique d'un état d'esprit, et si la musique est ‘*le langage des émotions*’ (Sundberg, 1998), alors il est un fait inéluctable que le *Haka* répond à plusieurs critères du chant, et doit être également considéré en tant que tel.

En guise de conclusion, le mot final qui servira de point d'orgue à cette étude, sera emprunté au poète qui, en termes raffinés et choisis, et sans doute bien mieux que nous n'aurions pu le faire, a su exprimer et résumer notre pensée :

« *All passionate language does of itself become musical with a finer music than the mere accent;  
the speech of a man even in zealous anger becomes a chant, a song...* »

Thomas Carlyle (1795-1881)

## Références bibliographiques

- ALLEN, G. D. (1975), Speech Rhythm: its relation to performance universals and articulatory timing, *Journal of Phonetics*, 3, p. 75-86.
- ARMSTRONG, A. (1964), *Maori games and Haka*, Reed.
- ASTESANO, C. (1999), *Rythme et discours : invariance et sources de variabilité des phénomènes accentuels en français*, Thèse d'université, Université de Provence, Aix-en-Provence.
- BIGGS, B. (1961), *The Structure of New Zealand Maori*, Archives of the Languages of the World, Bloomington, Indiana University.
- BIGGS, B. (1971), The Languages of Polynesia, *Current Trends in Linguistics*, vol. 8, p. 466-505, Mouton, The Hague.
- BOLINGER, D. (1985), *Intonation and its parts*, Edward Arnold.
- BOND, Z.S., MOORE, T.J. and GABLE, B. (1989), Acoustic-phonetic characteristics of speech produced in noise while wearing an oxygen mask, *J. Acoust. Soc. Am.*, 85 (2), p. 907-912.

- BOULET-GERCOURT, P. (2004), *La maison Bush*, Grasset, 262 p.
- BROWMAN, C., GOLDSTEIN, L. (1990), Tiers in articulatory phonology with some implications for casual speech, I. Kingston J. & Beckman M. (eds), *Papers in Laboratory Phonology*, Cambridge University Press, p. 341-376.
- CASTELLANOS, A. , BENEDI, J.M., CASACUBERTA, F. (1996), An analysis of general acoustic-phonetic features for Spanish speech produced with the Lombard effect, *Speech Communication*, vol. 20 (1), p. 23-35.
- CHAFCOULOFF, M., GUAÏTELLA, I., BOYER, J. (2001), Étude de la coordination temporelle des événements vocaux et gestuels dans l'exécution du Haka par les rugbymen néo-zélandais les All Blacks, in C. Cavé, I. Guaitella and S. Santi (eds), *Actes colloque Orage 2001*, L'Harmattan, p. 502-509.
- CHAFCOULOFF, M. (2003), A phonetic study of the Haka performed by the All Blacks, *Proceedings of the XVth International Congress of Phonetic Sciences*, Barcelona (Spain), 9-15<sup>th</sup> August, vol. 4, p. 2103-2106.
- DELATTRE, P.C. (1965), *Comparing the Phonetic Features of English, German, Spanish and French*, Julius Groos Verlag, Heidelberg.
- DELATTRE, P.C. (1969), First formant as a cue to place of articulation, *J. Acoust. Soc. Am.*, vol. 46 (1), p. 110.
- DI CRISTO, A. (1985), *De la microprosodie à l'intonosyntaxe*, Thèse de Doctorat d'État, Publications de l'Université de Provence, Aix-en-Provence.
- DUEZ, D. (1982), Silent pauses and non-silent pauses in three speech styles, *Language and Speech*, vol. 25 (1), p. 11-28.
- DURAND, P. (1985), *Variabilité acoustique et invariance en français : consonnes occlusives et voyelles*, Collection 'Sons et Parole', vol. 4, éditions du CNRS, 273 p.
- ESSEN, O. von (1957), *Allgemeine und angewandte Phonetik*, Berlin, Akademie Verlag.
- EZKENAZI, M. (1993), Trends in speaking styles research, *Proceedings of the 3<sup>rd</sup> European Conference on Speech Communication and Technology*, Berlin, vol. 1, p. 501-509.
- FANT, G. (1991), Units of temporal organization, stress groups versus syllables and words, *Proceedings of the 12<sup>th</sup> International Congress of Phonetic Sciences*, Aix-en-Provence, vol. 1, p. 247-250.
- FONAGY, I. (1983), *La vive voix : Essais de psycho-phonétique*, Payot, Paris.
- FÓNAGY, I. ; FÓNAGY, J. (1966) Sound pressure level and duration, *Phonetica*, 15, p. 14-21.
- FRAISSE, P. (1974), *Psychologie du rythme*, Presses Universitaires de France.
- GRAMMONT, M. (1933), *Traité de Phonétique*, Delagrave, 6<sup>e</sup> édition (1960), Paris.
- HARAWIRA, K.T.(1994), revised by Karetu, T., *Teach yourself Maori*, (3rd ed.), Reed.

- HARRIS, K.S. (1956), Some acoustic cues for the fricative consonants, *J. Acoust. Soc. Am.*, 28 (1), p. 160.
- HUSSON, R. (1960), *La voix chantée*, Gautier Villard, Paris.
- JASSEM, W. (1999), English Stress, Accent and Intonation revisited, *Speech and language Technologies*, vol. 3, Papers, Reports and technological Notes, Poznan, Polish Phonetic Association, p. 33-50.
- JESPERSEN, O. (1922), *Nature, évolution et origines du langage*, édition française (1976) 436 p., Payot, Paris.
- JUNQUA, J.C. (1992), The Lombard Reflex and its Role on Human Listeners and Automatic Speech Recognizers, *J. Acoust. Soc. Am.*, 93 (1), p. 510-524.
- JUNQUA, J.C. (1996), The influence of acoustics on speech production: a noise-induced stress phenomenon known as the Lombard reflex, *Speech Communication*, vol. 20 (1), p. 13-22.
- KARETU, T.S. (1993), *Haka! The dance of a noble People*, Reed.
- LABOV, W. (1964), Phonological correlates of social stratification, *American Anthropologist*, vol. 66, p. 164-176.
- LAFON, J.-C. (1961), *Message et phonétique. Introduction à l'étude acoustique et physiologique du phonème*, Presses Universitaires de France.
- LEHISTE, I. (1973), Rhythmic Units and Syntactic Units in Production and Perception, *J. Acoust. Soc. Am.*, 54 (5), p. 1228-1334.
- LOMBARD, E. (1911), Le signe de l'élévation de la voix, *Ann. maladies, oreilles, larynx, nez, pharynx*, 37, p. 101-119.
- MAC LAGAN, M., HARLOW R., KING J., KEEGAN P. and WATSON C. (2004a), New Zealand English influence on Maori pronunciation over time, *Te Reo*, vol. 47, p. 7-27.
- MAC LAGAN, M., HARLOW R., KING J., KEEGAN P. and WATSON C. (2004b), Acoustic analysis of Maori: Historical data, *Proceedings of the Australian Linguistic Society Conference 2004*, <http://dSPACE.library.syd.edu.au:8080/handle/123456789/105>.
- MORIER, H. (1961), *Dictionnaire de poétique et de rhétorique*, Presses Universitaires de France, première édition, 1210 p., Paris.
- O'SHAUGNESSY, D. (1982), A study of French Spectral Patterns for Synthesis, *J. of Phonetics*, 10 (4), p. 377-400.
- PETERFALVI, J.-M. (1970), Recherches expérimentales sur le symbolisme phonétique, *Monographies françaises de psychologie*, n° 19, Publications du CNRS, 174 p.
- PIKE, E., V. (1979), Word Stress and Sentence Stress in Various Tone Languages, *Proceedings of the 9<sup>th</sup> International Congress of Phonetic Sciences*, Copenhagen, vol. II, p. 410-416.
- RAMUS, F. *et al.* (1999), Correlates of Linguistic Rhythm in the Speech Signal, *Cognition*, 73 (3), p. 265-292.

- ROSSI, M. (1978), The Perception of non-repetitive Intensity Glides on Vowels, *J. of Phonetics*, 6 (1), p. 9-18.
- ROSTOLLAND, D. (1979), *Contribution à l'étude de l'audition de la parole en présence de bruit : caractéristiques physiques, structure phonétique et intelligibilité de la voix criée*, Thèse de Doctorat ès-sciences, Université Pierre et Marie Curie, Paris VI.
- ROUSSELOT, l'Abbé (1897), *Principes de phonétique expérimentale*, H. Welter, Paris.
- SCOTTO DI CARLO, N. (1982), *Cours d'initiation à la musique*, 61 p., Institut de Phonétique, Aix-en-Provence.
- SCOTTO DI CARLO, N. (1994), L'intelligibilité de la voix chantée, *Médecine des Arts*, n° 10, p. 1-16.
- SCOTTO DI CARLO, N., AUTESSERRE, D. (1992), L'organisation temporelle de la syllabe dans la parole et dans le chant, *Travaux de l'Institut de Phonétique d'Aix-en-Provence*, vol. 14, p. 189-204.
- SCOTTO DI CARLO, N., AUTESSERRE, D. (1993), Sons et sens dans l'opéra, *Médecine des Arts*, 4, p. 4-11.
- SCHULMAN, R. (1988), Articulatory Dynamics of Loud and Normal Speech, *J. Acoust. Soc. Am.*, 85 (1), p. 295-312.
- SOUND FORGE 4.0, *Mode d'emploi*, Waves system V 1.0, 1997.
- STANTON, B.J., JAMIESON, I. H., and ALLEN, G. (1989), Robust recognition of loud and Lombard Speech in the Fighter Cockpit Environment, *Proceedings of ICASSP-89*, p. 675-678.
- STATHOPOULOS, E.T., WEISMER, G. (1983), Closure duration of Stop Consonants, *J. of Phonetics*, 11 (4), p. 395-400.
- STEVENS, K.N. (1971), Airflow and Turbulence Noise for Fricative and Stop Consonants: Static Considerations, *J. Acoust. Soc. Am.*, 50, p. 1180-1192.
- STEVENS, K.N, HOUSE, A.S. (1963), Perturbations of vowel articulations by consonantal context: an acoustical study, *Journal of Speech and Hearing Research*, 6 (2), p. 111-128.
- STRAKA, G. (1964), L'évolution phonétique du latin au français sous l'effet de l'énergie et de la faiblesse articulatoires, *Travaux de Linguistique et de Littérature de Strasbourg*, vol. II, p. 17-98.
- SUNDBERG, J. (1974), Articulatory interpretation of the singing formant, *J. Acoust. Soc. Am.*, 55, p. 834-844.
- SUNDBERG, J. (1998), Expressivity in singing: A review of some recent investigation, *Log. Phon. Vocol.*, 23, p. 121-127.
- UMEDA, N.(1977), Consonant Duration in American English, *J. Acoust. Soc. Am.*, vol. 61, p. 816-858.
- VAN SUMMERS, W. *et al.* (1988), Effects of Noise on Speech Production: Acoustic and Perceptual Analyses, *J. Acoust. Soc. Am.*, 84 (3), p. 917-928.

VOSSIUS, G.J. (1597), *Commentariorum Rhetoricorum sive Oratoriorum Institutionum, Libri sex* (Opera III) Amsterdam.

WUNDT, W. (1886), *Eléments de psychologie physiologique*, trad. par E. Rouvier, Paris, Alcan.

ZWIRNER, E., ZWIRNER, K. (1937), Phonometrischer Beitrag zur Frage der Lesepausen, *Arch. Néerl. Phonet.* 13, p. 111-128.